

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match 3.4%; Score 99.5; DB 3; Length 544;  
Best Local Similarity 20.1%; Pred. No. 0.13;  
Matches 82; Conservative 47; Mismatches 115; Indels 163; Gaps 18;

QY		19	IVGAGLGSIVAYKLRKLRLNNAKIFEGAPDFGGVWNNRVPGARVDSETPYQLNIPEVW	78
Dd		50	VVGAGVSGLAAYRLOSNGYNVTVEADRADGG-----KIRTSBGGF-----YW	94
QY		79	KDMT-----WSC-----RYPOQKELLSVV-----97	
Dd		95	DEGANTMTGEWEASRLIDDLGLQDKOQYPNSQHVKRYIKVDGAPALIPSDPISLMKSSVL	154
QY		98	-----HHCDKTRGLRKDVYFEGAENVDAARYADLGTTWT	129
Dd		155	STSKIALTFEPFLKYKANTRNSGVKBSEBHSVSGVFCEGRHGREVVD--FYFDV--PFV	210
QY		130	VKTSAGHAVATAKYLILATGLLHKHTPALPGDANGKVITHSAMHEDFDAEQRVAVIG	189
Dd		211	AGTSAGD-----PESLRIHPALPNWLERSKYSVI-----VG	243
QY		190	AGATSQIVOELLAKKAQOVMTWMRRPSYCPLMRQR-----TMDRNBOTAKWAY--YPTLFE	243
Dd		244	A-----ILSKIAAKGD-----PVKTHDOSSCKRRNRVSFPHGOMOSLIN	284
QY		244	ASKRS-----RIGPFVQA-----PSVGIFEVSPQEORAYFEBELMERGAFNFL----	285
Dd		285	ALHNVEDDNVKLGTEVLISLACTFGCVPALGRWSISYDSKDSDXDLASNQTDFAVIMTA	344
QY		286	-ACQREVWMDKKANRLVIDEFWAQQ---TRSRIVNPARR-DIMAPLE	327
Dd		345	PLSNVRMKFTKGAPVYVDFLPKMYPILSIMYATPKQDDVKVPLE	391

Query Match 3.4% Score 99.5; DB 3; Length 544;  
Best Local Similarity 20.1%; Pred. No. 0.13;  
Matches 82; Conservative 47; Mismatches 163; Gaps 18;  
19 IVGAGLGSISAVYKURKLRNAKIFEGADPGGVHWRNYFGARVDSSTFYQNLNPEVM 78

Db 50 VVGAGVGLAAAYRLKQSGVNTVFEADRAGG-----KIRTNSEGGF-----VM 94  
Qy 79 KDMT-----WSC-----RYPDOKEKLLSYV----- 97  
Db 95 DEGANTWTEGEWEASRLIDDLGLQKQYVPSQHKRYIVKDGAPALIPSDPISLMKSSVL 154  
Qy 98 -----HHCOKIRGLRKQVYFGAEVVDARYARDLGTWT 129  
Db 155 STKSIALFFPEFLYKKANTRNSGVSEHLSSEVSGFCERHFGREVVD--YFVD--PFV 210  
Qy 130 VKTSAGHVATAKYLIILATGLLHKHTPALPLGLADFNKVIHSSAMHEDFDEAGORVAVIG 189  
Db 211 AGTSAGD-----PESLSIRHAFPALWNLERKYSVI-----VG 243  
Qy 190 AGATSIQIYOELAKKADQVTFMRRPSYCLPMRQR-----TMDRNEOTAKAY--YPTLFE 243  
Db 244 A-----ILSKLAAGD-----PVKTRHDSGSKRRNRVVSFSGHGMQSLIN 284  
Qy 244 ASRKS-----RIGFPVQA-----PSVGIFEVSPQREAYFEELMERGAFNEL--- 285  
Db 285 ALHNEVGDDNVKLGTEVLSLACTFDGVPALGRWSISVDSKDSGDKDLASNOTDAVIMTA 344  
Qy 286 -ACQYREVVDKANKRLVYDFWAKK---TRSRIVNPAKR-DLMAPLE 327  
Db 345 PLSNVRMKFTKGGAPVLDLFLPKMDYPLSLMVTAFKKDDVKKPLE 391

## RESULT 15

US-09-071-296-8  
; Sequence 8, Application US/09071296  
; Patent No. 6177245  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Eric R  
; APPLICANT: Volrath, Sandra  
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen  
; TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071.296  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/261,198  
; FILING DATE: 16-JUN-94  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elmer, James Scott  
; REGISTRATION NUMBER: 36,129  
; REFERENCE/DOCKET NUMBER: CGC 1748/CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-341-8614  
; TELEFAX: 919-341-8689  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 544 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-071-296-8

Query Match 3.4%; Score 99.5; DB 4; Length 544;  
Best Local Similarity 20.1%; Pred. No. 0.13;  
Matches 82; Conservative 47; Mismatches 115; Indels 163; Gaps 18;  
Qy 19 IVGAGLSGSAVYKRLKRLNAKIFEGAPDFGCVVHWNHRYPGARVDSSTPFYQNLNIPDV 78  
Db 50 VVGAGVGLAAAYRLKQSGVNTVFEADRAGG-----KIRTNSEGGF-----VM 94  
Qy 79 KDMT-----WSC-----RYPDOKEKLLSYV----- 97  
Db 95 DEGANTWTEGEWEASRLIDDLGLQKQYVPSQHKRYIVKDGAPALIPSDPISLMKSSVL 154  
Qy 98 -----HHCOKIRGLRKQVYFGAEVVDARYARDLGTWT 129  
Db 155 STKSIALFFPEFLYKKANTRNSGVSEHLSSEVSGFCERHFGREVVD--YFVD--PFV 210  
Qy 130 VKTSAGHVATAKYLIILATGLLHKHTPALPLGLADFNKVIHSSAMHEDFDEAGORVAVIG 189  
Db 211 AGTSAGD-----PESLSIRHAFPALWNLERKYSVI-----VG 243  
Qy 190 AGATSIQIYOELAKKADQVTFMRRPSYCLPMRQR-----TMDRNEOTAKAY--YPTLFE 243  
Db 244 A-----ILSKLAAGD-----PVKTRHDSGSKRRNRVVSFSGHGMQSLIN 284  
Qy 244 ASRKS-----RIGFPVQA-----PSVGIFEVSPQREAYFEELMERGAFNEL--- 285  
Db 285 ALHNEVGDDNVKLGTEVLSLACTFDGVPALGRWSISVDSKDSGDKDLASNOTDAVIMTA 344  
Qy 286 -ACQYREVVDKANKRLVYDFWAKK---TRSRIVNPAKR-DLMAPLE 327  
Db 345 PLSNVRMKFTKGGAPVLDLFLPKMDYPLSLMVTAFKKDDVKKPLE 391

Search completed: April 27, 2002, 07:46:43  
Job time: 6700 sec



---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 20:51:07 ; Search time 96.68 Seconds  
(without alignments)  
3112.349 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638

Sequence: 1 atgtcggccaccagcaactc.....tgcttcattgctccctgctaa 1638

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 139433 seqs, 91850399 residues

Total number of hits satisfying chosen parameters: 278866

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA-New\*

1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.8	2.5	177	5	US-09-975-254-7839
2	32.6	2.0	250	5	US-09-975-254-11803
3	32.6	2.0	1206	6	US-10-103-675-20
4	31.4	1.9	273	5	US-09-540-210B-7546
5	31.2	1.9	1365	5	US-09-540-209B-2547
6	31.1	1.9	2004	5	US-09-540-209B-2484
7	30.6	1.9	1532	5	US-09-974-702-1
8	30.4	1.9	241	5	US-09-975-254-25703
9	30.4	1.9	3855	6	US-10-107-431-182
10	30.4	1.9	45055	6	US-10-107-431-277
11	30.2	1.8	3188	1	PCT-US02-09921-464
12	30	1.8	32195	6	US-10-102-627-92
13	29.8	1.8	239	5	US-09-540-210B-31550
14	29.8	1.8	268	5	US-09-975-254-29303
15	29.6	1.8	39119	6	US-10-103-299-6650
16	29.6	1.8	1314	5	US-09-980-771A-6
17	29.6	1.8	1593	5	US-09-980-771A-8
18	29.6	1.8	1953	5	US-09-980-771A-4
19	29.6	1.8	2124	5	US-09-980-771A-2
20	29.6	1.8	3117	5	US-09-980-771A-1
21	29.6	1.8	10211	6	US-10-105-299-11932
22	29.6	1.8	20245	6	US-10-103-295-297
23	29.4	1.8	1600	6	US-10-114-104-206
24	29.4	1.8	4386	1	PCT-US02-07787-20
25	29.2	1.8	33249	6	US-10-105-299-10181
26	29	1.8	876	7	US-60-365-264-399

C	27	29	1.8	1218	5	US-09-540-209B-4669	Sequence 4669, Ap
	28	29	1.8	5043	6	US-10-116-802-17	Sequence 17, Appl
	29	29	1.8	19436	6	US-10-103-299-11327	Sequence 11327, A
C	30	29	1.8	22428	6	US-10-105-299-6872	Sequence 6872, Ap
	31	28.8	1.8	270	5	US-09-975-254-6508	Sequence 6508, Ap
C	32	28.8	1.8	2340	5	US-09-928-531-3	Sequence 3, Appl1
	33	28.8	1.8	2953	5	US-09-928-531-1	Sequence 1, Appl1
	34	28.8	1.8	15413	6	US-10-105-299-7826	Sequence 7826, Ap
	35	28.6	1.7	461	5	US-09-673-476-183	Sequence 183, App
C	36	28.6	1.7	714	5	US-09-540-209B-768	Sequence 768, App
	37	28.6	1.7	1077	5	US-09-540-209B-420	Sequence 420, App
	38	28.6	1.7	1278	5	US-09-540-209B-751	Sequence 751, App
	39	28.6	1.7	2292	6	US-09-540-209B-3275	Sequence 3275, Ap
	40	28.6	1.7	14083	6	US-10-105-299-9443	Sequence 9443, Ap
C	41	28.6	1.7	14105	6	US-10-105-299-9444	Sequence 9444, Ap
	42	28.4	1.7	308	6	US-10-015-219-226	Sequence 226, App
	43	28.4	1.7	317	6	US-10-099-926-254	Sequence 254, App
	44	28.4	1.7	322	6	US-10-099-926-755	Sequence 755, App
	45	28.4	1.7	322	6	US-10-099-926-841	Sequence 841, App

## ALIGNMENTS

RESULT 1  
US-09-975-254-7839  
; Sequence 7839, Application US/09975254  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Heck, Gregory R.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(15309)B  
; CURRENT APPLICATION NUMBER: US/09/97s,254  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US/09/263,191  
; PRIOR FILING DATE: 1999-03-05  
; NUMBER OF SEQ ID NOS: 31255  
; SEQ ID NO 7839  
; LENGTH: 177  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 700753780H1  
US-09-975-254-7839

Query Match 2.5%; Score 40.8; DB 5; Length 177;  
Best Local Similarity 55.7%; Pred. No. 0.0023;  
Matches 78; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
Qy 28 gattgttcgctgcagtcgagccatcatcgtaggagccgctcagcgcatctctgt 87  
Db 24 gattccctctcgtctcctcgtcatcgtaggagccgcatctcctgtagctgctg 83  
Qy 88 gtgtacaaatttggaagtctcagactcaacgcaaaattcttcgagggagcccccatttt 147  
Db 84 gcgaagtgcttgccgagaaacgtagagacctggtgatctctgagggcctcgaccgc 143  
Qy 148 ggcggcgctctggcactggaa 167  
Db 144 gtcggcgccagggatctgcaaa 163

RESULT 2  
US-09-975-254-11803  
; Sequence 11803, Application US/09975254  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Heck, Gregory R.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants

```
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 11803
; LENGTH: 250
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700905274H1
US-09-975-254-11803

Query Match      2.0%; Score 32.6; DB 5; Length 250;
Best Local Similarity 53.5%; Pred. No. 0.94;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 1 atgtcgccaccagcaactccagagcgaggttcttcgtcgcatcgagcgccatcgtt 60
Db 55 atctgctcatgagctcttcacgcgcgaagcgccgagtcagtgaacccgtcagagac 114
Qy 61 ggagccgcctccagcgccatctctgtgtacaaattgcaaaattgcagactcaacgccc 120
Db 115 gaagtcgacttcacatcgtctctgttacgaacaacctccggaagcccgctccaatgtcg 174
Qy 121 aaaatct 127
Db 175 gagaact 181

RESULT 3
US-10-103-675-20
; Sequence 20, Application US/10103675
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Nucleic
; FILE REFERENCE: 35800/207402
; CURRENT APPLICATION NUMBER: US/10/103,675
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US/09/749,278
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1206
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-103-675-20

Query Match      2.0%; Score 32.6; DB 6; Length 1206;
Best Local Similarity 53.5%; Pred. No. 1.9;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 624 aaccatgtttatgcgaagggccagctattgtctcccatcgccgcaacgcaacgatggatag 683
Db 982 accaaagattttggaaggagaagtacctggagtaccagcaggtggccaacagtgatcct 1041
Qy 684 gaacgaacagacagcctggaggcctactactacccacgctgtttgaacgagtcgaaagtc 743
Db 1042 gcaagatgatattcctctgtgggcccacgaagacccaagctgaacacgcaagatgaaagtc 1101
Qy 744 tcggatt 750
Db 1102 ctggagt 1108

RESULT 4
US-09-540-210B-7546
; Sequence 7546, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullany, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/595,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/303,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
```



```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/09/974,702
3 FILING DATE: 09-Oct-2001
4 CLASSIFICATION: <Unknown>
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: 09/017,981
7 FILING DATE: 03-FEB-1998
8 APPLICATION NUMBER: US60/037,854
9 FILING DATE: 07-FEB-1997
10 APPLICATION NUMBER: GB9705040.5
11 FILING DATE: 12-MAR-1997
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Hand, J. Mark
14 REGISTRATION NUMBER: 36,545
15 REFERENCE/DOCKET NUMBER: 19730
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: 732-594-3905
18 TELEFAX: 732-594-4720
19 TELEX: <Unknown>
20 INFORMATION FOR SEQ ID NO: 1:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 1532 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: double
25 TOPOLOGY: both
26 MOLECULE TYPE: cDNA
27 HYPOTHETICAL: NO
28 ANTI-SENSE: NO
29 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
30
31 US-09-974-702-1

```

	Query Watch	1.98;	Score 30.6;	DB 5;	Length 1532;
	Best Local Similarity	53.8;	Prod. No. 8.8;		
	Matches 63;	Conservative	Mismatches 54;	Indels 0;	Gaps 0;
QY	882	tgacaaaaggccaacgcactggtctctatgactctctgggccaaaagaactcgatctcgtat	941		
DB	1290	TGAGAGGCAGGCCAACTCTCTGGGCAAAATCTGGCCCTCCACAGAGGAGCCCTGGCAA	1349		
QY	942	cgtcaatccggcaaaagagagatctcaatggctcctctggagccgcctactgggttcgg	998		
DB	1350	CTTCTCTCAGTCCAGGCTTGAGCCACAGCCCTCCCGAGAGTCTCTCAGGTTTG	1406		

```

RESULT      8
US-09-975-254-25703
; Sequence 25703, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: plants
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 25703
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700965057H1
US-09-975-254-25703

```

Query Match 1.9% Score 30.4; DB 5; Length 241;  
Best Local Similarity 50.7%; Pred. No. 4.5;  
Matches 73; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```

Db 43 gctcgacacattctactacgtcagctctcgggatacagcgtcggcggctcccggtgcg 102
      ||||| ||| || || |||| ||| ||| ||| ||| ||| |||
Qy 507 tcattcgagtcctgcgcagaaagacttcgacgcagagggccagagagtcgcgcgtcatcgg 566
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103 cggcctctcgctcgtctcttcgcctcgacgcagcgggaaacggcggtcatcatcga 162
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 567 tgcggggggccacaagcatccagat 590
      ||||| ||| ||| ||| ||| |||
Db 163 ctcgggcacctccgtgaccgggt 186
      ||||| ||||| ||||| ||||| |||||

RESULT 9
US-10-107-431-182
; Sequence 182, Application US/10107431
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING AND DISTINGUISHING
; TITLE OF INVENTION: BIOSYNTHETIC LOCI
; FILE REFERENCE: 3001-705
; CURRENT APPLICATION NUMBER: US/10/107,431
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 182
; LENGTH: 3855
; TYPE: DNA
; ORGANISM: Streptomyces mobaraensis
US-10-107-431-182

```

	Query Match	1.9%;	Score 30.4;	DB 6;	Length 3855;
	Best Local Similarity	49.4%;	Pred. No.15;		
	Matches	79;	Conservative	0;	Mismatches 81;
					Indels 0;
					Gaps
Qy	446	tgtctccacaggaagcacactccgcgactccgcggactccgcatttccaacggaaggtag	505		
Db	1190	tgtgtgacgycgaccccaaccgcagatcgactggcagcgccacgccatcccgctcggccca	124		
Qy	506	ttcatttcgagtgcctggcagcagaatttcgacgagagggccacagagtcgcgcgtcatcg	565		
Db	1250	ccggggccgtgccttgcccgagcgcgcaccccgccgcgcgtctccgcgcttcg	130		
Qy	566	gtgccggggccacagcatccagatgtttcaggagttgagc	605		
Db	1310	qctacagcgcacgttgccacacgtctctctgaqagcagc	1349		

```

RESULT 10
US-10-107-431-277
; Sequence 277, Application US/10107431
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING AND DISTINGUISHING O
; TITLE OF INVENTION: BIOSYNTHETIC LOCI
; FILE REFERENCE: 3001-70S
; CURRENT APPLICATION NUMBER: US/10/107.431
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 45055
; TYPE: DNA
; ORGANISM: Streptomyces mobaraensis
US-10-107-431-277

Query Match 1.9%; Score 30.4; DB 6; Length 45055;
Best Local Similarity 49.4%; Pred. No. 46;

```







```

RESULT 15
US-10-105-299-6650
: Sequence 6650, Application US/10105299
: GENERAL INFORMATION:
: APPLICANT: Rosen, et. al
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: PS950
: CURRENT APPLICATION NUMBER: US/10/105,299
: CURRENT FILING DATE: 2002-03-26
: NUMBER OF SEQ ID NOS: 15197
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 6650
: LENGTH: 39119
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-105-299-6650

Query Match      1.8%; Score 29.8; DB 6; Length 39119;
Best Local Similarity 52.9%; Pred. No. 66;
Matches 64; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 1195 atggccttgaacacacacgagtgagcctgaaggaggtgtggaagatggcatatct 1254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21242 ataggcctaaaaaaagacaggggtaagtagcagtggtggaagagggacgcacct 21301

Qy 1255 acttatatgggagtgcttctctcatgcttccccaatgccttcttcgtcgccacggctcaa 1314
    | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21302 agatcatgaatacatatcatatgactttctcatatataccttaactcccaagtgcc 21361

Qy 1315 g 1315
Db 21362 g 21362

Search completed: April 27, 2002, 01:35:52
Job time: 17065 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 19:22:16 ; Search time 5001.19 Seconds  
(without alignments)  
4420.550 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638

Sequence: 1 atgtcgccaccgcaactc.....tggtcatgctccctgctaa 1638

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pla:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.4	5.8	544	9 AW180687	AW180687 Mga0835f
2	88.8	5.4	496	12 AQ962221	AQ962221 LERGE08TF
3	88	5.4	621	9 AW745217	AW745217 LGL_387_F
4	77.8	4.7	393	9 AT005115	AT005115 AT005115
5	69.4	4.2	596	12 AQ967995	AQ967995 LERIN94TF
6	69	4.2	421	12 AQ967996	AQ967996 LERIN94TF
7	67	4.1	398	9 AW784433	AW784433 d5912a1.r
8	52.8	3.2	529	10 BI949834	BI949834 HVSME1001
9	51.8	3.2	498	10 BG507355	BG507355 sac58a01.
10	49.2	3.0	546	10 BE760936	BE760936 an_0959 A
11	48	2.9	472	10 BI700307	BI700307 sag66c04.
12	46	2.8	915	10 BF065971	BF065971 HV_CEB001
13	45.8	2.8	421	9 BE188259	BE188259 PSI3583 I
14	45.6	2.8	539	9 AA556804	AA556804 646 loblo
15	44.6	2.7	396	10 BG157908	BG157908 EM1_6_D05
16	44.6	2.7	450	10 BF177238	BF177238 EM1_1_D01
17	44.6	2.7	551	10 BG464165	BG464165 EM1_70_D1

18	44.6	2.7	557	10 BG463588	BG463588 EM1_50_D1
19	44.6	2.7	560	10 BG36283	BG36283 EM1_23_F0
20	43.4	2.6	589	10 BF263256	BF263256 HV_Cra000
21	43	2.6	510	10 BG158097	BG158097 EM1_8_D10
22	41.6	2.5	693	12 AQ259083	AQ259083 ndxb00221
23	41.4	2.5	191	10 BG411699	BG411699 EM1_60_HO
24	41.2	2.5	545	10 BI498406	BI498406 sail13f05.
25	40.4	2.5	602	10 BG454500	BG454500 NF099E03L
26	39.8	2.4	585	10 BG523525	BG523525 32-7 Stev
27	39.6	2.4	382	10 C97940	C97940 C97940 Rice
28	39.4	2.4	541	12 AQ991295	AQ991295 Rfc02194
29	39.4	2.4	1101	12 CNS01793	AL108460 Drosophil
30	39	2.4	422	12 AZ131793	AZ131793 OSJNB011
31	38.2	2.3	628	10 BG526271	BG526271 58-43 Ste
32	37.8	2.3	619	9 AW267708	AW267708 EST305836
33	37.8	2.3	649	12 BH489445	BH489445 BGO849TR
34	37.6	2.3	464	12 PR0030900	AL027269 Fugu_rubr
35	37.2	2.3	363	9 AW981361	AW981361 EST392514
36	37.2	2.3	519	10 BI174155	BI174155 OSTF010H1
37	37.2	2.3	532	10 BJ116593	BJ116593 BJ116593
38	37.2	2.3	561	9 AU208788	AU208788 AU208788
39	37.2	2.3	565	9 AU198737	AU198737 AU198737
40	37.2	2.3	570	10 BJ107504	BJ107504 BJ107504
41	37.2	2.3	700	10 BJ150341	BJ150341 BJ150341
42	37	2.3	460	10 C98651	C98651 C98651 Rice
43	37	2.3	503	10 BM134375	BM134375 WHE0489_D
44	36.8	2.2	504	10 BE705434	BE705434 SC02_1290
45	36.8	2.2	535	9 AV637658	AV637658 AV637658

## ALIGNMENTS

RESULT AW180687  
LOCUS Mga0835f Mga Library Mycosphaerella graminicola linear EST 17-NOV-1999  
DEFINITION 5' similar to cyclohexanone monooxygenase, mRNA sequence.

ACCESSION AW180687  
VERSION AW180687.1 GI:6447882  
KEYWORDS EST.

SOURCE Mycosphaerella graminicola.

ORGANISM Mycosphaerella graminicola

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et

Chaetothryomycetes incertae sedis; Mycosphaerellaceae;

Mycosphaerella

1 (bases 1 to 544)

Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.

A group of expressed cDNA sequences from the wheat fungal leaf

blotch pathogen, Mycosphaerella graminicola (Septoria tritici)

Fungal Genet. Biol. 29 (2), 118-133 (2000)

20374020

Contact: Hargreaves JA

Cell Biology Department

IACR-Long Ashton Research Station

Long Ashton, Bristol, BS41 9AF, UK

Tel: +44(0)1275 392181

Fax: +44(0)1275 394281

Email: john.hargreaves@bbsrc.ac.uk

Insert Length: 1000 Std Error: 0.00

Seq primer: M13 reverse.

Location/Qualifiers

1..544

/organism="Mycosphaerella graminicola"

/strain="Strit"

/db\_xref="taxon:54734"

/clone="Mga0835"

/clone\_lib="Mga Library"

/note="Vector: pSPORn1; Library constructed from cultures

utilizing ammonium ions as a source of nitrogen"

150 a 157 c 132 g 105 t

BASE COUNT

ORIGIN

Query Match									
Best Local Similarity 5.8%; Score 95.4; DB 9; Length 544;									
Matches 268; Conservative 0; Mismatches 266; Indels 3; Gaps 1;									
QY	795	ccccgagcagcgggagccctatttcgaagagttgtgagagcgtggggcccttaattttct	854						
DB	11	CCAGAAGCAGAGAGACAGCTTCTATGAAAAGCTATGGCAAGAGGGCGGTTCGGATAC	70						
QY	855	tgcttgcagtagccgagagtcattgtgacaaaaagccacccgactggtctatgactt	914						
DB	71	GCTCAGCAGCTACAAAGGACATGCTCTTCGACACAGAGGCCAATATGAGGCGCTACAA	130						
QY	915	ctgggcccacaaagactcgtatctcgtatctcaatccggcgaagagatctcatggctcc	974						
DB	131	CTGGAGGAAAAGGTCCTCGCTCGCTCAAGGATGAGGAGAACGAGAAATCTTGCACC	190						
QY	975	tctggagccgcgctactggttgcgtaccagcgctcccccactggagagcgactactaca	1034						
DB	191	AAAGTCCACCTCACCCATGGGGAACAAGCGCCCACTCTCGAGCAGCGCTTCTACAA	250						
QY	1035	aatctgtgacaaagccgagcgtcgaaattgtgaatttagaacaatcgccattgtgctgt	1094						
DB	251	GGTTGTGCGACCCCGACACAGTCAACATCATCGACGTCAACGAGAAATCCCATCGAAATGT	310						
QY	1095	tacaaagacaggtgtgctctttagtgacggcgagcgaagggaatcgacacgatcgtgct	1154						
DB	311	---CAAGCGCGAGGGTCTTGTCAACCCGAAAGGCGCTCGAGGAGATTGATATTGATCT	367						
QY	1155	ggcgaggggttttcgacagtttcaactggtcattgacacatatgggttgaaacaaaga	1214						
DB	368	TGCTACAGGCTTCGATTCGGTCACTGTGGTTCGCTTGGCCAGCTCAACATTTCAAGACAACAA	427						
QY	1215	cggagcgacctgaaagggtgtgaaagatggcatatctattatgaggagctctctc	1274						
DB	428	GGGTCAACCATTTGCCAACCATCGAAGCAGCGATGCAAAACATCCATGGGAATTCGCAT	487						
QY	1275	tcatggttccccaatgccttcttctgctccacggtcctaaagcccgacctctcttc	1331						
DB	488	TCCTGGCTTCCCGCAACATGTTCTTCTGCTACGCTCCACAGGCACCTACTGCTTCTC	544						
RESULT 2									
A0962221 A0962221 496 bp DNA linear GSS 28-JAN-2000									
LOCUS LERGE08TF LERG Arabidopsis thaliana genomic clone LERGE08, DNA									
DEFINITION sequence.									
ACCESSION A0962221									
VERSION A0962221.1 GI:6789922									
KEYWORDS GSS.									
SOURCE thale cress.									
ORGANISM Arabidopsis thaliana									
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;									
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.									
1 (bases 1 to 496)									
Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,									
Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.									
Genomic survey sequencing of Landsberg erecta ecotype of									
Arabidopsis thaliana and identification of sequence-based									
polymorphisms									
JOURNAL Unpublished (2000)									
COMMENT Contact: Xiaoying Lin									
The Institute for Genomic Research									
9712 Medical Center Dr., Rockville, MD 20850, USA									
Tel: 301 838 0200									
Fax: 301 838 0208									
Email: at@tigr.org									
For additional information, see http://www.tigr.org/tdb/at/at.html									
Seq primer: TF									
Class: Shotgun.									
Location/Qualifiers									
1. .496									
FEATURES									
source									
1									
-									
-									

/organism="Arabidopsis thaliana"									
/strain="Landsberg erecta"									
/db_xref="taxon:3702"									
/clone="LERGE08"									
/clone.lib="LERG"									
/note="Organ: Leaf; Vector: pUC19JK; Total genomic DNA was									
sheared to 0.4-0.7 kbp before ligation."									
BASE COUNT 106 a 149 c 146 g 95 t									
ORIGIN									
Query Match 5.4%; Score 88.8; DB 12; Length 496;									
Best Local Similarity 50.2%; Pred. No. 5e-14;									
Matches 247; Conservative 0; Mismatches 242; Indels 3; Gaps 1;									
QY	225	cgaagtgtgaaagactggacctggtcttgcgcgtatctcctgacacgaagaagtgctgc	284						
DB	4	CGAGTTGCAGTAGGAAGTGGCACTGGCAGCAACGCTACGCGCGGACGGAGATTTTGC	63						
QY	285	atatattccacctgtgacaagatccggggcttgaaaaagacgtctacttcggagctga	344						
DB	64	CTACGCCAACACAGTTGGCGATCGCTTCAACTCGCGGGCGGATATCCAGTTTCGATAC	123						
QY	345	ggtggtgtgatgcgcggtatgccagagatcttggcaacctggagctgtcaagaagctcg	404						
DB	124	AGTCGATCGCGCGGTTCGATGAAAACGCCCAATGCTTCGTGGTGCAGTCTCGGAC	183						
QY	405	ccatgtgcgcagcgaagatctctattctcgtactcgggtgtgtctccacaggaagcac	464						
DB	184	AAAGATCTGCACGGCGCAATTCGTCGCTGGCTACCGGCTGCCCTCCACGCCACGTAC	243						
QY	465	tcccgacctccccggcctgcgcgatttcaacggggaaggtgatcttcattcgagtcgt	521						
DB	244	GCGGACTTCAAGGGGACTCTCGGATTTCAAGGGCAGCGCTTATCACACGGACACT	303						
QY	522	gcacgaagacttcgacgcagagggccagagtcgcgcgtatcgtgcccggggccacaag	581						
DB	304	GCACGAGCGGTAGATTTACCGCCCAAGCGGCTGCGCGTGATTTGGCAGCGGATCG	363						
QY	582	catcagattgttcaggagttggccaaagaaggtgcacaggttaacctgtttatgcgaag	641						
DB	364	GATCAGACGATTCGGTGTGATCGGTGAACAAGCCAGCGAGCTTTTCGTATTTCAC	423						
QY	642	gcgcgactattctgcgccatcgggcgaacgaacgatgtaggaacgaacagacagctg	701						
DB	424	CCGGAATTCAGTGTCCCGCGCGCAATGCCGNACTGACGAAGGAGAACCCAGCG	483						
QY	702	gaaggcctacta 713							
DB	484	TCGGGACAACTA 495							
RESULT 3									
A0962217 A0962217 621 bp mRNA linear EST 19-JUL-2000									
LOCUS LG1_387_F06.bl_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA									
DEFINITION sequence.									
ACCESSION A0962217									
VERSION A0962217.1 GI:7658955									
KEYWORDS EST.									
SOURCE sorghum.									
ORGANISM Sorghum bicolor									
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC									
clade; Panicoideae; Andropogoneae; Sorghum.									
1 (bases 1 to 621)									
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C. and Pratt, L.H.									
An EST database from Sorghum: Light-grown seedlings									
Unpublished (2000)									
JOURNAL									
COMMENT Contact: Cordonnier-Pratt MM									
Department of Botany									
The University of Georgia									
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA									

/organism="Arabidopsis thaliana"  
/strain="Landsberg erecta"  
/db\_xref="taxon:3702"  
/clone="LERGE08"  
/clone\_lib="LERG"

BASE COUNT 106 a 149 c 146 g 95 t  
ORIGIN  
/note="Organ: Leaf; Vector: pUC19JK; Total genomic DNA was  
sheared to 0.4-0.7 Kbp before ligation."

Query Match 5.4%; Score 88.8; DB 12; Length 496;  
Best Local Similarity 50.2%; Pred. No. 5e-14;  
Matches 247; Conservative 0; Mismatches 242; Indels 3; Gaps 1;

QY	225	cgaagatggaagactggacctgcttgcgcgtatctctgaccagaaagagttgctgtc	284						
DB	4	CGAGTTGCAGTAGGAATGGGACTGCGAGCGAACGCTACGCGCGCGAGGACATTTGCG	63						
QY	285	atatgttccactgtgacaaagatccggggttgagaaaaagacgtctactctcggagctga	344						
DB	64	CTACGCCAACACACGTTCGCGATCGCTTCAACCTGGCGGCGGATATCCAGTTGATACCG	123						
QY	345	ggtggttgatgcgcgtatgccagagatctggccctggactgtcaagacgtcggctgg	404						
DB	124	AGTCGATCGCGCGGTTCGATGAAACGCCAATGCTTGGTGGTGACGTCGTCGACGG	183						
QY	405	coatgttgcgaaggcaaaagtattctatctcgtcaggggttgctccacagaaagacac	464						
DB	184	AAAGATCTCTACGGCGCAATTCGTGCTACCGGCTGCCCTTCCAAACGCGACGTAC	243						
QY	465	tcctgcactcccgccctgcgcgatttcaacgggaggtgattcattcagtgacctg---	521						
DB	244	GCCGACTTCAAGGACACTCTCGGATTTCAGGGCAGCGTTTATCACACGGACACTGG	303						
QY	522	gcacgaagacttcagcagaggccagagagtcacctcctcgtgcgcggggccacaag	581						
DB	304	GCACGAGCGGTAGATTTCACCGCAAGCGCTGCCCGTGTGGCACCAGGATGTCAGC	363						
QY	582	catccagatgttcaggagttggccaagaaggtgacaggttaaccatgtttatcgaa	641						
DB	364	GATCCAGACGATTTCGCTGAACCAAGCCAGCCAGCTTTTCGTATTTCACAGCGAC	423						
QY	642	gccagactatgtctgcccctcgtgcgaacgacgatgtaggaacgacagacagacctg	701						
DB	424	CGCAATTTCAGTGTCCCGCGCGCAATGCCGAAGGAGGAGCAACGCCAGCGGTT	483						
QY	702	gaagcctacta 713							
DB	484	TCGGGCAACATA 495							

RESULT 3  
A0745217 A0745217 621 bp mRNA linear EST 19-JUL-2000  
LOCUS LG1\_387\_F06.b1\_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA  
DEFINITION sequence.  
ACCESSION A0745217  
VERSION A0745217.1 GI:7658955  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 621)  
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C. and Pratt, L.H.  
An EST database from Sorghum: light-grown seedlings  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860  
Fax: 706 542 1805

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV

High quality sequence stop: 604

POLYA-No.

Location/Qualifiers

1. .621

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_lib="Light Grown 1 (LG1)"

/note="Organ: 10- to 14-day-old light-grown (greenhouse)

seedlings; Vector: Lambda Zap; Site:1: XhoI; Site:2: EcoRI

; The library was made from poly-A RNA in the cloning

vector lambda Zap II. Clones to be sequenced were

prepared by mass excision."

164 a 156 c 142 g 138 t 1 others

BASE COUNT  
ORIGIN

Query Match 5.4%; Score 88; DB 9; Length 621;

Best Local Similarity 50.2%; Pred. No. 9,6e-14;

Matches 276; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

QY 41 catgcgaacatcatcgttggagccggcctcagcgggaatctctgtgtacaaattgc 100

Db 63 CACTTGATGCTATTGTTGGAGCTGGATCTCCGGAATCAGCGCTCTGTATCAACTTC 122

QY 101 gaaactcagactcaacgcgcaaatcttcgagggagcccccggatttggcggtctggc 160

Db 123 GACAAACAGGGCTAAATGTGAAGTTGTTTCAAAAGGCTCTCTGAAGTTGGCGCTTGGT 182

QY 161 actggaacgctacccctggcgctcgtgttgattcggagacgcccctctaccacactgaac- 219

Db 183 ATTGGAACCAATACCAGATGACGACCGCATTCCTCCGAGCACTCTACAGATTAGCT 242

QY 220 --attccgaagtatgaaagactggagactgtgtctgttcgctatctcagcagaagaagt 277

Db 243 GGGATAAAGACAACTGCAGACTTACCATGGCCAAACATGTTCTTCCCAAGAGAAA 302

QY 278 tgcgtcatatgttcaaccactgtgacaagatccggggcttgagaaagacgtctacttcg 337

Db 303 TTCAAGGCTATCTAAAGCAGCTGCTGGACCACTATGACTTGAGGCAGCATATTACATTTT 362

QY 338 gagctgagtggttgatgcgcgatgcccagagatcctggccacttgactctcaagacgt 397

Db 363 CTCGGAGCTCAAGTCTGCGTCATTT-----GATCAAGTGTATGGACTGTCAATTTT 416

QY 398 cggctggccatgttgcgacgggcaagaagtattctattctcgtacggggttgctccacagga 457

Db 417 CTACTGGGACAGCTTCAAGACACATCTTGATCTCTGCTGTGGGTACATGTCAACAAC 476

QY 458 agcacactcccgactcccgctcgcgatttcaacggggaagtgatttcattcagtg 517

Db 477 CCGTTGGCCCTACCTCTCTGGCCAGGACAAAGTTCAAGGGCGCAACAGTATCACACAGGAC 536

QY 518 cctgcaacgaagacttcagcagagagggccagagatcgccctacgttcggtccgggcca 577

Db 537 TTTGGCTGAGCTGCTGATCGATGCGGAGCAAGCGAGTCGGAATCATTTGGCACCGGCTCTT 596

QY 578 caagcatcca 587

Db 597 CAGGGCGTCA 606

RESULT 4

AT005115

LOCUS

DEFINITION

AT005115 Pleurotus ostreatus cDNA clone MFB1-A03, mRNA

393 bp mRNA linear EST 21-MAR-2001

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AT005115  
AT005115.1 GI:13419973

oyster mushroom.

Pleurotus ostreatus

Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

Agaricales; Pleurotaceae; Pleurotus.

1 (bases 1 to 593)

Lee,S., Kim,K., Lee,K., Kim,K., Yoo,Y. and Lee,C.

Expressed sequences in fruiting body of Pleurotus ostreatus

Unpublished (2001)

Contact: Beom-Gi Kim

Division of applied microbiology

Institute of Agricultural Science and Technology (NIAS)

249 Seodundong Kwonseonku, Suwon 441707, Korea

Tel: 82-331-290-0347

Fax: 82-331-290-0399

Email: bkimyes@da.go.kr

Submitted through BRIC(Biological Research Information Center) of

Korea

URL: http://bric.postech.ac.kr/

GeneNuri No. KS105130.

Location/Qualifiers

1. .393

/organism="Pleurotus ostreatus"

/cultivar="ASI 2029"

/db\_xref="taxon:5122"

/clone="MFB1-A03"

/clone\_lib="POMFBO1"

/dev\_stage="mature fruiting body"

/lab\_host="E.coli"

/note="Vector: lambda TriPEX2; Site:1: SfiIA; Site:2:

SfiIBI; average insert size:1500 bp; initial pfu:5 \* 10 7;

Isolation of total RNA from the mature fruiting body

cultivated in poplar tree sawdust bottle"

109 a 85 c 107 g 92 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 193; Conservative

0; Mismatches 192; Indels

0; Gaps

0;

4.7%; Score 77.8; DB 9; Length 393;

50.1%; Pred. No. 5.7e-11;

0; Mismatches 192; Indels

0; Gaps

0;

QY 1064 tgaatcagaacaatcgcccatgttgctgttacaagacagagtgctgttgagtgacg 1123

Db 1 TGGACCTCAACGAAAGCCCACTTGACGAGATTACATCAACCGGGGTTCGTACCAAGATG 60

QY 1124 gcagcaagagggaatgcgacacgacgtcgtcgcgacgggtttccgacagtttcaactggct 1183

Db 61 GAACAGAAATATAATGTGGACCTGTGGTCATGGCCACGGGTTTCGATATGGGAACGTGG 120

QY 1184 cattgacacatatgggcttgaaaaaacaagcagcagagtgagacctgaaggaagtggaag 1243

Db 121 GCTTTAAGGACATGTGAATCGTAGGGACAAACGGGGCGCTTTTCACGATTAATGGGCCA 180

QY 1244 atggcatatctattatgggagtcctctctcctcgtggtcttcccaatgccttcttcgtcg 1303

Db 181 ATGGGGTGAAGTCGTATCTCTGGCATGCTAGGCTCTGGGTTTCCAAACATGTTTATGGTAT 240

QY 1304 ccacggctcaagcccgccgacgtcccttccacggcccaacgacatcatagaaccacagtcg 1363

Db 241 ACGGCCCTCAGCAGCAGCGCGCTTTACCAATGGCCCTTACATGTGCTGATTCACAGTTG 300

QY 1364 acttgatcgcgatacaaatgtcaasgttggggcgagcagcagcagtcgcttgagggca 1423

Db 301 ATTGGGTAAACCAATTGCATTGAATATATGACGAAGAAATTCGCTCGCCCGCATCGAAGCA 360

QY 1424 cgaatcagcacaagagagcagtcg 1448

Db 361 GTCAAAAAGCGGAACCTGGATTGGAC 385

RESULT 5





```

RESULT 9
BG507355      498 bp      mRNA      linear      EST 28-NOV-2001
LOCUS      sac58a01.y1 Gm-cl062 Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION      ID: Gm-cl062-3866 5' similar to TR:Q9SVQ1 Q9SVQ1 HYPOTHETICAL 46.5
                KD PROTEIN. ;, mRNA sequence.
ACCESSION      BG507355
VERSION
KEYWORDS
SOURCE
ORGANISM
    Glycine max
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
REFERENCE
    1 (bases 1 to 498)
    Shoemaker,R., Kaim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna
    ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
    Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
    ,Y., Persson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
    ,R., Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
    ,R., Waterston,R. and Wilson,R.
    Public Soybean EST Project
    Unpublished (1999)
    Contact: Shoemaker R/Public Soybean EST Project
    Public Soybean EST Project
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: est@watson.wustl.edu
    This clone is available through: resgen, Invitrogen Corp, 2130
    South Memorial Parkway Huntville, AL 35801 For further information
    call: (800)-533-4363 or contact via email: ccufesgen.com
    High quality sequence stop: 415.
FEATURES
    source
    1..498
        /organism="Glycine max"
        /db_xref="taxon:3847"
        /clone="GENOME SYSTEMS CLONE ID: Gm-cl062-3866"
        /clone_lib="Gm-cl062"
        /tissue_type="stem tissue of greenhouse grown plants"
        /dev_stage="1 month old"
        /lab_host="DH108"
        /note="vector: pBluescript II SK+; Site.1: EcoRI; Site.2:
        XhoI; The cDNA library was constructed from mRNA isolated
        from stem tissue of 1 month old greenhouse grown plants
        for the cultivar Raiden. Complementary DNA was
        synthesized from mRNA using a primer consisting of a
        poly(GT) sequence with a XhoI restriction site. EcoRI
        adapters were ligated to the blunt-ended cDNA fragments
        followed by XhoI digestion. The cDNA fragments were
        directionally cloned into the EcoRI-XhoI restriction site
        of the pBluescript vector. The ligated cDNA fragments were
        transformed into DH108 host cells (GibcoBRL). This library
        was constructed in the laboratory of Dr. Randy Shoemaker."
BASE COUNT      126 a 108 c 131 g 131 t
ORIGIN
Query Match      3.2%; Score 51.8; DB 10; Length 498;
Best Local Similarity 50.4%; Pred. No. 0.0015;
Matches 125; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY 416 cggcaagatctatctcgtcaggggtgtgtccacaggaagcacactccgcactcc 475
    |||||
DB 242 CTGGAAGTACCTGCTGTCGCCAGTGGCGAAGCCGCGAGCTCTCTGCTCAGATTC 301
    |||||
QY 476 ccggctcgcgatttcacggaggtgatttcatttcgatttcgacgagacttcg 535
    |||||
DB 302 AAGCTTGGAGATTTCAATCGAAGAGTATTCTACTGCTACAAAAACGGGACG 361
    |||||
QY 536 acgcagagagggccagagagtcgcgcctcattcgtgtgccggggccacacagacttcatttc 595
    |||||

```

---

```

Db 362 AGTTTAAGAAATAAATGTTCTGTTGTCGGATCTGCGCAATTCGCGCATGAGATTGCGC 421
    |||||
QY 596 agaggttgcccaagaagcgtcaccaggttaaccatgtttatcggaagccgagctattgtc 655
    |||||
DB 422 TAGACTTGTGCAATTTTGTGTGCCAAACCTTCATCATTTGTCGACGCCGCTTCATTNTC 481
    |||||
QY 656 tgcccatg 663
    |||||
DB 482 TTTCGAGG 489
    |||||
RESULT 10
BE760936      546 bp      mRNA      linear      EST 24-OCT-2000
LOCUS      an_0959 Aspergillus niger, pYES2 (XhoI-EcoRI) Aspergillus niger
DEFINITION      CDNA clone 0959 3', mRNA sequence.
ACCESSION      BE760936
VERSION
KEYWORDS
SOURCE
ORGANISM
    Aspergillus niger.
    Aspergillus niger.
    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
    Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE
    1 (bases 1 to 546)
    Tsang,A. and Storms,R.
    Aspergillus niger Expressed Sequence Tags
    Unpublished (2000)
    Contact: Tsang, A.
    Dept. of Biology
    Concordia University
    1455 demaisonneuve Blvd. West, Montreal, Quebec, CANADA, H3G 1M8
    Tel: 514-848-3405
    Fax: 514-848-4504
    Email: tsangevax2.concordia.ca
    PCR PRIMERS
    BACKWARD: GCGTGCAATGTAGCGTGACATAAC
    Insert Length: 546 Std Error: 0.00
    POLYA=No.
FEATURES
    Location/Qualifiers
    1..546
        /organism="Aspergillus niger"
        /db_xref="taxon:5061"
        /clone="0959"
        /clone_lib="Aspergillus niger, pYES2 (XhoI-EcoRI)"
        /lab_host="E. coli"
        /note="vector: pYES2; Site.1: Xho-I; Site.2: EcoRI; cDNA
        was synthesized with ZAP kit (Stratagene) using poly(A)+
        RNA isolated from Aspergillus niger grown in both complete
        and minimal media. Synthesis was primed with oligo(dT)
        primer/XhoI-linker. EcoRI adapters were later ligated to
        polished ends. EcoRI-XhoI-digested cDNA was ligated with
        EcoRI-XhoI-digested pYES2 (Invitrogen Corp). This vector
        permits expression of our library in yeast."
BASE COUNT      106 a 140 c 165 g 127 t
ORIGIN
Query Match      3.0%; Score 49.2; DB 10; Length 546;
Best Local Similarity 50.5%; Pred. No. 0.0085;
Matches 145; Conservative 0; Mismatches 139; Indels 3; Gaps 1;
QY 321 aaaaagcgtctacttcgagctgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 380
    |||||
DB 527 AAAGGATTCTCTTCTTAATACGGTTGTAGTATCAGCGGAATACGATCTCTCTGACG 468
    |||||
QY 381 ctggactgtcgaagcgtcggtcgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 440
    |||||
DB 467 ATGGCAATCCGACAGCAATGCTGCTGTCACGCATCCGCGANATACCTTCATCTCTGAC 408
    |||||
QY 441 ggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 500
    |||||
DB 407 CGGATTCCGCCCAAGCCACACGTCCTCTGATGCGCAGGATGACACCTTCCTGCGTGGAG 348
    |||||

```



QY 501 ggtgattcattcagtgctggcagcaagact---tcgacgcagagggccagagagtcgc 557  
 DB 347 AATCCACCACTCTGCTCTGCTGGCCGAGACGGCTCGACGTGCAAGCAACGCTCGC 288  
 QY 558 cgtcatcggtgcggagccagcaagactccagattgttcagagtgagg 604  
 DB 287 CGTATCGGCACGGCGGCTCGCGCATCTCACTCACTCAAGCCCTGG 241

## RESULT 11

BI700307  
 LOCUS  
 DEFINITION  
 ID: Gm-cl082-1280 5' similar to TR:D9SVU0 Q9SVU0 HYPOTHETICAL 48.1  
 KD PROTEIN. ;, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 soybean.

Query Match 2.9%; Score 48; DB 10; Length 472;  
 Best Local Similarity 49.6%; Pred. No. 0.017;  
 Matches 123; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

## REFERENCE

AUTHORS  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)

## TITLE

JOURNAL  
 COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

Trace considered overall poor quality. This clone is available through: Resgen, Invitrogen Corp. 2150 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 1.

## FEATURES

source

1. .472  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl082-1280"  
 /clone\_lib="Gm-cl082"  
 /tissue\_type="Roots of bulked Minsoy x Noir RI progeny"  
 /lab\_host="BHL08"  
 /note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2: XhoI; The mRNA was isolated from roots of bulked Minsoy x Noir RI progeny that are phenotypically susceptible to flooding damage. The plants were flooded (submerged) for 24 hours prior to harvest. Dr. Tara VanToal generously provided the isolated mRNA. The roots were flash-frozen in liquid nitrogen. StrataGene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGACTAGTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-

DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University.\*

BASE COUNT 127 a 77 c 147 g 121 t  
 ORIGIN

QY 416 cggcaaatatctcattctcgtcacgggtgtctccacagggaagcacactccgcactcc 475  
 DB 54 CTGGAAGTACTCTGACGTGCCAGTGGGAAACCGCCGCAAGCTGCTCTGCTCAAAATTC 113  
 QY 476 cggcctcgcgatttcaacgggaagtgtattcattcagtgctgcgcagcagactcgc 535  
 DB 114 AAGCTTGGAGAGAATCAATGGAAGAGTGAATCAAGTCTTTTAAACGGGAACG 173  
 QY 536 acgcagagggccagagagtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 595  
 DB 174 AGTTTAAAGAAATAACATGTTCTGTGTTGCGAATCTGGCAATTCGGCATGGAGATTCGC 233  
 QY 596 aggagttggccaaagaagcgtgaccaggtaccatgtttatgcgaaggccgcagctattgtc 655  
 DB 234 TAGACTTGTCCAAATTTTGTGTCACAACTTACATCATTTGTCGACGCCGGGTTATTTTC 293  
 QY 656 Tgcccattg 663  
 DB 294 TTTCCAGG 301

## RESULT 12

BF065971  
 LOCUS  
 DEFINITION  
 HV\_CEB0014H22f Hordeum vulgare seedling green leaf EST library  
 HV\_CEB0014H22f, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae; Hordeum.  
 1 (bases 1 to 915)  
 Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Oates, R. and Main, D.

Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected incompatible (M156) seedling leaf cDNA library  
 Unpublished (2001)  
 On Oct 17, 2000 this sequence version replaced gi:10842610.  
 Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu

## JOURNAL

## COMMENT

Total hq bases = 551

Seq primer: AATTAACCCCTCACTAAAGG

High quality sequence stop: 642.

Location/Qualifiers

## FEATURES

source

1..915

/organism="Hordeum vulgare"

/cultivar="C116151 (M1a6)"

/db\_xref="taxon:4513"

/clone="HV\_CEB0014H22f"

/clone\_lib="Hordeum vulgare seedling green leaf EST

library HVCDNA0005 (Blumiera challenged)"

/tissue\_type="seedling green leaf"

/lab\_host="SOLR"

/note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;

C.I. 16151 (M1a6) plants were greenhouse grown in the R

wise lab at Iowa State University, Ames, IA; 7 day old

green seedlings were challenged with isolate 5874 (AvrM1a6

) of Blumeria graminis f. sp. hordei, and leaves were

harvested 20 and 24 hr post-inoculation and snap frozen;

uninoculated leaves were harvested 20 hr post-inoculation

(Wei, Wise). In the TJ Close lab at the University of

California, Riverside, total RNA was prepared from each

sample pool, equal quantities of all three RNA pools were

combined, poly(A) RNA was purified from the mixture, one

primary unamplified cDNA library was made, and 1 million

pfu were in vivo excised to give phagescript SK(-) cDNA

phagemids (Choi, Close). Phagemids were plated and picked

at the Clemson University Genomics Institute (CUGI) (Begum

, Palmer, Frisch, Atkins and Wing). Plasmid DNA

preparations, DNA sequencing and sequence analysis were

performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates

, Rambo, Main). The sequence has been trimmed to remove

vector sequence and contains a minimum of 100 bases of

phred value 20 or above. For more details on library

preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order

this clone see <http://www.genome.clemson.edu/orders> Also

see Close TJ, Wing R, Kleinhofs A, Wise R (2001)

Genetically and physically anchored EST resources for

barley genomics. Barley Genetics Newsletter 31:29-30.

(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT  
ORIGIN

152 a 317 c 304 g 142 t

Query Match 2.8%; Score 46; DB 10; Length 915;

Best Local Similarity 50.5%; Pred. No. 0.092;

Matches 112; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 417 gcaaaagtatctctctgctacgggggtgtctccacaggaagacacactccgcactccc 476

Db 395 GCGGAGGTACCTGGTGGTGGCCACGGGGAGTTCGATGAGAAGGTGTCGCCGAGGTGCC 454

QY 477 cgccctcgccgattccacggagaggtgattcattcagtgccctggcagacgaacttca 536

Db 455 TGGGCTGGACAGTTCCTCCGCGCAGGCCATCCATCGCGGGAGTACCGGTCCGCGAGGG 514

QY 537 cgcagagggccagagatgcgcgtatcggtgcgcggggggccacgaagcatcagattgtca 596

Db 515 GATGCGGGGAAGGAAGTCTCGTGTGCTGCGGTGCGGCAATTCGGAATGGAGATCGCCCT 574

QY 597 ggaattggcccaagaaggtgaccaggtaaccattgtttatg 638

Db 575 TGACCTGGCCCAAGCGCGCGCGCTCCATCTGTCGTCG 616

RESULT 13  
LOCUS

BE188259 421 bp mRNA linear EST 25-SEP-2000

DEFINITION PS135R3 In vitro expressed cDNAs Cladosporium fulvum cDNA, mRNA

sequence.

ACCESSION BE188259

VERSION BE188259.1 GI:8667498

## KEYWORDS

SOURCE

ORGANISM

Cladosporium fulvum.

Cladosporium fulvum

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et

Chaetothyrionomycetes incertae sedis; Mycosphaerellaceae; mitosporic

Myosphaerellaceae; Cladosporium.

REFERENCE

1 (bases 1 to 421)

AUTHORS

Segers, G.C., Henricot, B., Rasmussen, S.W. and Oliver, R.P.

TITLE

In vitro expressed genes of Cladosporium fulvum (Segers, Henricot,

Rasmussen, Oliver)

JOURNAL

Unpublished (2000)

COMMENT

Contact: R.P.Oliver

Murdoch University

SABC, Perth 6150, Western Australia

Tel: +61-8-9360-7404

Fax: +61-8-9360-6303

Email: roliver@central.murdoch.edu.au

High quality sequence stop: 421

POLYA-No.

Location/Qualifiers

1..421

/organism="Cladosporium fulvum"

/strain="Race 4"

/db\_xref="taxon:5499"

/clone\_lib="In vitro expressed cDNAs"

/tissue\_type="Mycelium"

/note="Vector: Lambda bluescript; A mixture of mycelial

cultures grown in liquid B5 for 48 hours and transferred

for 24 hours to media lacking carbon, nitrogen or

supplemented with hydrogen peroxide"

BASE COUNT 107 a 108 c 116 g 86 t 4 others

ORIGIN

Query Match 2.8%; Score 45.8; DB 9; Length 421;

Best Local Similarity 56.8%; Pred. No. 0.068;

Matches 83; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 45 cgacgcacatctgtggagccgctcagcgcatctctgtgtacaaattgcgaaa 104

Db 174 CGATGTCCTCATTTGCGGTGGTGTTCACGGCGCCATGAGCTTGACAGGCAAGGAA 233

QY 105 gctcagactcaacgcaaaattctcagggagccccccgatttgcgcgtctgcactg 164

Db 234 GAATGGCTTCAACCCGTGCTCTACGAAGCTGGTCTCGGATTCGGCGGTACATGGCGATG 293

QY 165 gaaccgctaccctggcgctcgtgttg 190

Db 294 GAACATCTATCCGGGANCCTCGGTTG 319

RESULT 14

AA556804

LOCUS

646 Loblolly pine C Pinus taeda cDNA clone 6C11G, mRNA linear EST 28-AUG-1998

ACCESSION AA556804

VERSION AA556804.1 GI:3365818

KEYWORDS

EST.

SOURCE

loblolly pine.

Pinus taeda

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE

1 (bases 1 to 539)

AUTHORS

Allona, I., Quinn, M., Shoop, K., Swope, K., St.Cyr, S., Carls, J.,

Riedl, J., Retzel, E., Campbell, M.M., Sederoff, R. and Whetten, R.W.

TITLE

Analysis of xylem formation in pine by cDNA sequencing

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 95 (16), 9693-9698 (1998)

MEDLINE

98356220

COMMENT

Contact: Ross Whetten

Forest Biotechnology Group

North Carolina State University

Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 05:30:19 ; Search time 104.96 Seconds  
(without alignments)  
576.746 Million cell updates/sec

Title: US-09-882-694A-3

Perfect score: 2902

Sequence: 1 MGATGNSRGDCSVACDAIIV.....QICQEKVANWGDGFDVLHAPC 545

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_032802.\*
- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
  - 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
  - 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
  - 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
  - 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
  - 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
  - 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
  - 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
  - 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
  - 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
  - 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
  - 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
  - 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
  - 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
  - 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
  - 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
  - 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
  - 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
  - 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
  - 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
  - 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
  - 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	545	22	AY72634
2	481.5	16.6	603	22	AB85324
3	262.5	9.0	558	21	AA80087
4	213.5	7.4	535	19	AAW49699
5	210.5	7.3	471	22	AA39176
6	208.5	7.2	550	22	AA40962
7	199.5	6.9	532	19	AAW49700
8	199.5	6.9	594	22	AAW79793
9	192.5	6.6	558	22	AAW78809
10	184	6.3	434	22	AAG92928
11	170.5	5.9	532	17	AA97549
					Exophiala spinifera
					R. ruber cyclodode
					Flavin monooxygena
					Human flavin-conta
					Human polypeptide
					Human flavin-conta
					Human protein SEQ
					Human protein SEQ
					C glutamicum prote
					Human flavin-conta

12	163.5	5.6	286	22	AAB93890	Human protein sequ
13	161.5	5.6	95	21	AA02090	Cyclohexanone mono
14	161	5.5	568	22	ABG21986	Novel human diagno
15	156.5	5.4	416	22	ABBS9569	Drosophila melanog
16	153	5.3	219	22	AA79223	Corynebacterium gl
17	151	5.2	450	22	AAU33718	Pseudomonas aerugi
18	150.5	5.2	342	22	AAG91411	C glutamicum prote
19	148.5	5.1	390	22	AAG92804	C glutamicum prote
20	146.5	5.0	429	22	ABBS9688	Drosophila melanog
21	145.5	5.0	342	22	AAU60184	Protonibacterium
22	145.5	5.0	452	21	AAG40334	Arabidopsis thalia
23	143.5	4.9	342	22	AAG80036	Pseudomonas putida
24	143.5	4.9	342	22	AAG80036	Pseudomonas putida
25	142.5	4.9	452	21	AAU13800	Arabidopsis thalia
26	142.5	4.9	459	21	AAU13799	Arabidopsis thalia
27	137.5	4.7	461	21	AAG46666	H. pylori cytoplas
28	131.5	4.5	326	18	AAW21022	S. aureus thioredo
29	125	4.3	312	22	AAU00834	Staphylococcus aur
30	124	4.3	311	20	AAU29854	Cyclohexanone mono
31	123.5	4.3	88	21	AA02093	Putative P. abyssal
32	122	4.2	345	22	ABBS96222	Glutamine-oxogluta
33	122	4.2	506	20	AAU01313	C glutamicum prote
34	122	4.2	506	22	AAG89954	N. asteroides stra
35	120.5	4.2	455	20	AAU29930	Cyclohexanone mono
36	118	4.1	92	21	AA02092	Novel human diagno
37	117.5	4.0	1396	22	ABG29766	Cyclohexanone mono
38	116	4.0	73	21	AA02094	Cyclohexanone mono
39	116	4.0	570	21	AAU67416	Arabidopsis E3-1 s
40	116	4.0	623	21	AAG39829	Arabidopsis thalia
41	116	4.0	638	21	AAG39828	Arabidopsis thalia
42	116	4.0	658	21	AAG39827	Arabidopsis thalia
43	115.5	4.0	509	15	AAU47583	NADH oxidase. Amp
44	113.5	3.9	310	20	AAU29855	Staphylococcus epi
45	113.5	3.9	310	22	AAG82034	S. epidermidis ope

ALIGNMENTS

RESULT 1  
AAU72634  
ID AAU72634 standard; Protein: 545 AA.

AC AAU72634;

DT 02-MAY-2001 (first entry)

DE Exophiala spinifera flavin monooxygenase, a fumonisin degradative enzyme.  
Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;  
detoxification; mycotoxin; animal feed; human feed; sludge;  
transgenic plant; transgenic animal; microbial spray.

OS Exophiala spinifera.

PN WO200105980-A1.

XX 25-JAN-2001.

XX 14-JUL-1999; 99WO-US15824.

XX 14-JUL-1999; 99WO-US15824.

XX (PION-) PIONEER HI-BRED INT INC.

XX (CURA-) CURAGEN CORP.

XX Duwick JP, Madcox J, Gilliam J, Folkerts O, Crasta OR;

XX WPI; 2001-147345/15.

XX N-PSDB; AAD02690, AAD02691.

XX Novel polynucleotides encoding Exophiala degradative or transport  
PT









CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 471 AA;

Query Match 7.3%; Score 210.5; DB 22; Length 471;  
Best Local Similarity 22.6%; Pred. No. 1.7e-11;  
Matches 89; Conservative 61; Mismatches 160; Indels 83; Gaps 13;  
QY 19 IVGAGLGSISAVYKLRKLRNAKIFGAPDFGVWNNRYPGARVDSETPFYQLNPEVW 78  
Db 7 VIGAGVGLISIKCCVDEGLPTCFERTEDIGGVW---rfkenvedgrasiyqsvvntns 63  
QY 79 KDWTFSCRYPDQKELLSYVHCDKIRGLR-----KDYVFGAEVVDARYARDL---G 126  
Db 64 KEMSCFSDFPMPDFPFIHNSKLLLEYFRIFAKKFDLLKYLQFTTVLSVRKCPDFSSG 123  
QY 127 TWVTKT--SAGHVATAKY--LILATGLLHRKHPT--ALPGLADFNKGVIHSSAWHEDFDAE 181  
Db 124 QKVVVTQSGNGKQSAVDAVMVCSGHHILPHILKSPGMEFKQYFHSRQYKHPDGF 183  
QY 182 GORVAVIGAGATSIQIQVELAKKADQVTWFMRRPSYCLPMRQRTMDRNEQTAWKAYPTL 241  
Db 184 GKRIIVGMNGNSGDAVELSKNAQVFISTRHGTWV----- 221  
QY 242 FEASRKRISRGFPVQAPSGVIFEVSPQREAYPEELWGERGAFNPLACQYREVVDKANKRL 301  
Db 222 ---sriedgyp-----wdsvftrfmslrnvlprtavkwmieqgmnr- 262  
QY 302 VYDFWAKKTRSRIVNPAKRLMAPLEPPYWFQTKRSPLESYIEMLDKPSVEIVNLEQSP 361  
Db 263 ---wfnhen-----Yglepqnkylmkepvinddvpsrllcgakv---kst 302  
QY 362 IVAVTKTGVLISDGSKRE-CDTIVLATGFDSEPT 393  
Db 303 VKELTETSAIFDGTVEENIDVIFATGY-SFS 334

RESULT 6  
ID AAM40962 standard; Protein; 550 AA.  
XX  
AC AAM40962;  
DT  
XX 22-OCT-2001 (first entry)  
DE Human polypeptide SEQ ID NO 5893.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX

PD 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US34263.  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0582317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao JA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
DR N-PSDB; AAI60118.  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX Example 2; SEQ ID NO 5893; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX Sequence 550 AA;

Query Match 7.2%; Score 208.5; DB 22; Length 550;  
Best Local Similarity 22.6%; Pred. No. 3.3e-11;  
Matches 89; Conservative 60; Mismatches 161; Indels 83; Gaps 13;  
QY 19 IVGAGLGSISAVYKLRKLRNAKIFGAPDFGVWNNRYPGARVDSETPFYQLNPEVW 78  
Db 22 VIGAGVGLISIKCCVDEGLPTCFERTEDIGGVW---rfkenvedgrasiyqsvvntns 78  
QY 79 KDWTFSCRYPDQKELLSYVHCDKIRGLR-----KDYVFGAEVVDARYARDL---G 126  
Db 79 KEMSCFSDFPMPDFPFIHNSKLLLEYFRIFAKKFDLLKYLQFTTVLSVRKCPDFSSG 138  
QY 127 TWVTKT--SAGHVATAKY--LILATGLLHRKHPT--ALPGLADFNKGVIHSSAWHEDFDAE 181  
Db 124 QKVVVTQSGNGKQSAVDAVMVCSGHHILPHILKSPGMEFKQYFHSRQYKHPDGF 198  
QY 182 GORVAVIGAGATSIQIQVELAKKADQVTWFMRRPSYCLPMRQRTMDRNEQTAWKAYPTL 241  
Db 184 GKRIIVGMNGNSGDAVELSKNAQVFISTRHGTWV----- 236  
QY 242 FEASRKRISRGFPVQAPSGVIFEVSPQREAYPEELWGERGAFNPLACQYREVVDKANKRL 301  
Db 237 ---sriedgyp-----wdsvftrfmslrnvlprtavkwmieqgmnr- 277  
QY 302 VYDFWAKKTRSRIVNPAKRLMAPLEPPYWFQTKRSPLESYIEMLDKPSVEIVNLEQSP 361  
XX





Matches 81; Conservative 59; Mismatches 139; Indels 137; Gaps 12;

QY 19 IVGAGLSGISAVYKLRKLRNAKIFEGAPDGGVWHNRYPGARVDSETPFYQLNIPETW 78  
DB 7 VIAGVSGISIKCCVDEDLEPTCERSDDIGGLW---KTESSKDGMTVYKSLVINCVC 63  
QY 79 KD-----WTWSCRYPDQKELLSYVHHCCKIRGLRKDVYFGAEV 116  
DB 64 KEMSCYDFPHEDYPMFMNHKFDYIGFAEHFDILKVIQFTTVCSTKRPDFS--- 120  
QY 117 VDARYARDLCTWTYKTA---GRVATAYKYLILATGLLRKHHTP--ALPGLADFNKGVHIS 171  
DB 121 -----ETGQDQVLTETEGKQNAVFDAMVCTGHHFNHPLPLEAFGPHKFKGILHS 173  
QY 172 SAWHEDFADGECORVAVIGAGATSIQIVQELAKKADQVTMFMRPSVCLPMRQRTMDRNEQ 231  
DB 174 QEYKIPGFGKCVLVIGIGTGGDIAGELSTAAQVLLSTRGTGTVL----- 221  
QY 232 TAWKAYPTLFEASRKSRIQFPVQAPSVGIFEVSPQREAYFEELWERGAFNPLACQYRE 291  
DB 222 -----GRSDWGYP-----YNNM----- 234  
QY 292 VWOKKANRLYDFWAKKTRSRIVNPAKRDLMAPLEPPYFGTKRSPLESDDYE---M 346  
DB 235 -----VTRCCSFIAGVPSFLN-----WIGERKLNKRFNHEDVGLSIT 274  
QY 347 LDKPSVEIVNLE-----QSPIVAVTKTVGLLSDGSKRE-CDTIVLATGF 389  
DB 275 K9KKAKFLVNDLPNCILGALTKTSVLEFETSATSAVFDGTVEENIDVVFTTGY 330

RESULT 10  
AAG92928  
ID AAG92928 standard; Protein: 434 AA.  
AC AAG92928;  
DT 26-SEP-2001 (first entry)  
DE C glutamicum protein fragment SEQ ID NO: 6682.  
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;  
XX organic acid synthesis.  
OS Corynebacterium glutamicum.  
PN EP1108790-A2.  
PD 20-JUN-2001.  
PF 18-DEC-2000; 2000EP-0127688.  
PR 16-DEC-1999; 99JP-0377484.  
PR 07-APR-2000; 2000JP-0159162.  
PR 03-AUG-2000; 2000JP-0280986.  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI: 2001-376931/40.  
DR N-PSDB: AAG68147.  
XX Novel polynucleotides derived from Corynebacterium, for identifying  
PT mutation point of a gene, measuring expression of a gene, analysing  
PT expression profile of pattern of a gene and identifying homologous gene  
PT  
XX  
PS Claim 17; SEQ ID NO: 6682; 246pp + Sequence Listing; English.  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Corynebacterium glutamicum. These

are useful for identifying the mutation point of a gene derived from a mutant of corynebacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from corynebacterium, and identifying a homologue of a gene derived from corynebacterium. Corynebacterium bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention.  
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office.

XX Sequence 434 AA;

Query Match 6.3%; Score 184; DB 22; Length 434;  
Best Local Similarity 21.2%; Pred. No. 5.6e-05;  
Matches 96; Conservative 68; Mismatches 184; Indels 124; Gaps 18;

QY 18 IVGAGLSGISAVYKLRKLRNAKIFEGAPDGGVWHNRYPGARVDSETPFYQLNIPETW 76  
DB 10 IIGGGGGGIGALGARLKRIGVPALIDKASRPDQWR-SFY-----HSLCHDP 57  
QY 77 WVKDWTWSCRYPD-----QKELLSYVHHCCKIRGLRKDVYFGAEVVDARYARDLGTWT 129  
DB 58 WYDHLPIYIPFDHPVFTPKDKMGDWLHYVGLMDL--DYWTNTECLRASYNEDTKQWD 115  
QY 130 V---KTSAGHVATAYKYLILATGLLRKHHTP--ALPGLADFNKGVHISAWHEDFADAE-QGRV 185  
DB 116 VTVRDGAEATLPTQLVMTATGMSGSPNKPPTLPQGDKFQGEIRHSAE-HPGGVDVDRKNV 174  
QY 186 AVIGAGATSIQIVQELAKKADQVTMFMRPSVCLPMRQRTMDRNEQTAWKAYPTLFEAS 245  
DB 175 VVIGANNSHNDICADLYSNGAKPMIQRSTHIV-----TSDSLMRCVFGP----- 220  
QY 246 RKSRIQFPVQAPSVGIFEVSPQREAYFEELWERGAFNPLACQYREVWVWKKANRLYDFP 305  
DB 221 -----LYSADAVEAG-----IDTDCADLIFAS 242  
QY 306 WAKTRSRIVNPA-----KRDLMAPLEPPYW---FQTKRSPL-----ESDYVEMLD 348  
DB 243 WPYKVLPGVQKQAFDKIREDDKEFYDKLENAGFLIDFGDDSGILKYLIRGSGYY--ID 300  
QY 349 KPSVEIVNLEQSPIVA-----VTKTVGLLSDGSKRECDTIVLATGF---DSFTSGSLTHM 399  
DB 301 VGASLVADGKIPVRSNVSLEDVKNSSVLTGDTLPEDVIVLATGYGMMNNVWVQLVDQ 360  
QY 400 GLKNKHGV-----DLKEVWK 414  
DB 361 ETADKVGPCVGLSGTETKDPGPWEGELRUMWK 392

RESULT 11  
AAR97549  
ID AAR97549 standard; Protein: 532 AA.  
AC AAR97549;  
DT 14-SEP-1996 (first entry)  
DE Human flavin-containing monooxygenase.  
KW Monooxygenase; flavin; xenobiotic; carcinogen.  
XX Homo sapiens.  
OS Homo sapiens.  
PW EP12932-A2.  
XX 22-MAY-1996.  
PD 17-NOV-1995; 95EP-0118149.  
PF 18-NOV-1994; 94JP-0284902.



DB 185 gkrvliigsggdlaiveisqtakqvfistrqawil----- 222  
 QY 242 FEARKEKRIQFPQAVSVEIFVSPQREAYFELNKGAFNFIACQYREVWVDKKN-R 300  
 DB 223 --hrvdygypa-----dvlfsrlthf--lw-kicggslanky-----lekkingr 264  
 QY 301 LVYDFWAKKTRSRIV 315  
 DB 265 fdhemfglqpkhryv 279

RESULT 13  
 ID AAB02090  
 AC AAB02090  
 DT 03-JAN-2001 (first entry)  
 DE Cyclohexanone monooxygenase sequence.  
 KW Identification; prokaryote; polymerase chain reaction; PCR;  
 KW amplification; primer; differential display;  
 KW picric acid degradation; gene cluster; open reading frame; ORF;  
 KW dehydratase; dehydrogenase; transcription factor;  
 KW Acyl-CoA synthase; NADPH oxidoreductase;  
 KW cyclohexanone monooxygenase.  
 OS Unknown.  
 QX WO2000049177-A2.  
 PN 24-AUG-2000.  
 PD 17-FEB-2000; 2000WO-US03989.  
 PF 19-FEB-1999; 99US-0120702.  
 PR (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA Rouviere P;  
 PI WPI; 2000-587069/55.  
 DR Differential display method using a large number of arbitrary primers  
 PT for RT-PCR used to isolate novel differentially expressed prokaryotic  
 PT genes  
 XX Example 8; Page 62; 56pp; English.

CC A new method to identify differentially expressed prokaryotic genes  
 CC using a large number of arbitrarily primed polymerase chain  
 CC reactions comprises separating two populations of microbial cells,  
 CC where a first population is contacted with a stimulating agent;  
 CC extracting total RNA from both microbial cell populations;  
 CC amplifying the extracted RNA from both populations by preparing a  
 CC collection of at least thirty-two different arbitrary primers, where  
 CC each primer comprises a common and a variable region; individually  
 CC contacting each primer of with a sample of extracted RNA from the  
 CC two populations under conditions where two sets of amplification  
 CC products are produced; purifying the two sets of amplification  
 CC products; identifying the amplification products generated in the  
 CC first population which differ from products generated from the  
 CC second population as differentially expressed genes; and optionally  
 CC sequencing the identified differentially expressed genes. The  
 CC advantage over previous methods is that previous methods of  
 CC differential display to clone genes using thirty-two or thirty  
 CC primers have isolated four and one genes, respectively. The  
 CC new method using a greater number of primers has isolated  
 CC twenty-one induced gene fragments.  
 XX Sequence 95 AA:

Query Match 5.6%; Score 161.5; DB 21; Length 95;  
 Best Local Similarity 37.0%; Pred. No. 8.7e-08;  
 Matches 34; Conservative 17; Mismatches 40; Indels 1; Gaps 1;  
 QY 450 IETQVLDIADTIARKEAEHATSVEATKSAQEAWSIMIAKMHETFLPLTDSWMTGGINPG 509  
 DB 3 letqvewlsdtvayaeerneleiraleptpeaeewtqtcdianatlftrgdswifganvpg 62  
 QY 510 KATRALTFIGGIALYEQICQKVA-NWDGFDV 540  
 DB 63 kkpavlfyglgnyrnviagvvadsyrgfel 94

RESULT 14  
 ID ABG21986  
 AC ABG21986  
 DT 18-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #21977.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 OS Homo sapiens.  
 QX WO2000175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US08631.  
 PF 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 PA (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang VT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS86173.  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity

PS Claim 20: SEQ ID No 52345; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 568 AA:

Query Match 5.5%; Score 161; DB 22; Length 568;  
Best Local Similarity 19.0%; Pred. No. 1.5e-06;  
Matches 106; Conservative 73; Mismatches 208; Indels 172; Gaps 20;

QY 19 IVGAGLSGISAVYKLRKLNKAKIFEGAPDGGVWHNNRYPGA--RVDSETPFY---QLN 78  
DB 27 ivgagsglasikccleegleptcfersddlggfwfthv---eegrasykxsvnsnc 83  
QY 79 KDWTCSCRPDOKE-----LLSYVHCDKIRGLRKDVYFGAEVVDARYARD- 124  
DB 84 kemcsysdfpfedynypvpsqfleylkmyanhfd---llkhqfktkvcsvtkcsds 139  
QY 125 --LGTWTWTKT---SAGHVATARYLILATGLLHRKHTP--ALPGLADFNKGVIHSAWHE 177  
DB 140 avsggqevvthmheekgesaifdvmvctgfltnpyipldsfpginafkqyfhsgqykhp 199  
QY 178 FDEGQVAVIGAGATSIQIVQELAKKADQVTFMRRPSYCLPMRQTMDRNEQTAWKAY 237  
DB 200 difkdrvligmngsgtdiaveashaekvfistgggwi---srifdsg--ypwdmv 254  
QY 238 YPTLFASRSKRIGPVGQAPSVGIFEVSPQREAYFEELWERCANFL-----A 286  
DB 255 imtrfgmlrnsipcp-----vcwmerkinmwlnhanyglipedr 296  
QY 287 QVYRE-VMDKXANRLVYDFAKTRGRIVNPAKRDLMAPLEPPYWGTRKSPLESYYE 345  
DB 297 tqikevlnelpgrii-----tgkvfirps----- 322  
QY 346 MLDKPSVEIVNLEOSPPIVAVTKTGVLLSDGSKRE-CDTIVLATGDFDSTGSLTHMLKNK 404  
DB 323 -----ikevksvifnntskcepidliivfatgytfafpfidesvvkve 366  
QY 405 HGVDLKEVWKDGIYSTVMGVSFGHPNAPFAFFVATAQAPTV-----LSNGPTII 450  
DB 367 dg-----qaslykyfip-----ahlgkptlaighlhlplglxplgtetqv 408  
QY 451 ETQVDLIADTIKLEAHEA--TSVENYKSAQEAW-----SIIA 487  
DB 409 rwavsgsxkvitrlpppsvmeelnarkenkpswfglcycalqsdityidelltyint 468  
QY 488 KMNEHTLFLPLTD---SWWT 503  
DB 469 kplfsilwtdptsqglwt 487

## RESULT 15

ID ABB59569 standard; Protein; 416 AA.

AC ABB59569;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 5499.

XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX N-PSDB; ABL03672.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX Disclosure; SEQ ID NO 5499; 2lpp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins  
XX (ABB57737-ABB72072).  
XX The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 416 AA;

Query Match 5.4%; Score 156.5; DB 22; Length 416;  
Best Local Similarity 19.2%; Pred. No. 2.5e-06;  
Matches 102; Conservative 67; Mismatches 155; Indels 207; Gaps 21;

QY 19 IVGAGLSGISAVYKLRKLNKAKIFEGAPDGGVWHNNRYPGA--RVDSETPFY---QLN 73  
DB 6 ligagtaglccarhsiangfettvfelsdriggtwvyneatgavngldvhsmyknlrtn 65  
QY 74 IP-----EVMKDWTCSCRYPDOKELLS-VYHCDKIRGLRKDVYFGAEVVDARYAR 123  
DB 66 lpkevngfpdfeiganeasyvrsdeicdfingyanhfd---lkkhikfidsyvi-rvllq 119  
QY 124 DLGTWTWV--KTSAGHVATARY---LILATGLLHRKHTPALPGLADFNKGVIHSSAWHEDF 178  
DB 120 rktkwqylfkdlvtnkietqyfdkvivanghynpcnysqipmmerfkqgflshbfrsre 179  
QY 179 DAEGQVAVIGAGATSIQIVQELAKKADQVTFMRRPSYCLPMRQTMDRNEQTAWKAY 238  
DB 180 vfeqksvligagpsgmdlsnlsrtadrvti---shhltidigqhsffenvq----- 228  
QY 239 PTLFEASRSKRIGFPVQAPSVGIFEVSPQREAYFEELWERCANFLACQYREVWVDKKA 298  
DB 229 -----qkpdv-----reldekaf----- 242  
QY 299 NRLVYDEWAKKTRSRIVNPAKRDLMAPLEPPYWGTRKSPLESYYEMLDKPSVEIVNLE 358  
DB 243 ----- 242  
QY 359 QSPIVAVTKGVLLSDGSKRECDTIVLATGDFDSTGSLTHMLKNKHGVD-----L 409  
DB 243 -----fvdsyqfcdtvcfctgykyafpfit-----vdsyiyvedayv 280  
QY 410 KEVWKDGI---STYMGVFSHGFPNAPFAFFVATA-----QAPTVLS--NGPTIETQVDLIAD 459  
DB 281 qelykqcinrpsmail--glp--fyvcaaqmmdlgarfimsyngsnelpstedmikd 336  
QY 460 TIAKLEAERHATSVEATKSAQEAWSIIMAKMNEHTLFP-----LTD5WMTGG 505  
DB 337 trdrm-----gkwaeglrkrhahmlgpkdydftdlsqtag 373

Search completed: April 27, 2002, 07:45:31  
Job time: 8112 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 19:27:16 ; Search time 8172.36 Seconds  
(without alignments)  
4336.319 Million cell updates/sec

Title: US-09-882-694A-2  
Perfect score: 1638  
Sequence: 1 atctgcgccaccagcaactc.....tgcttcattgctccctgctaa 1638

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Maio:\*  
1: /cgn2\_6/ptodata/2/pna/US06005\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06006\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US06007\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1638	100.0	1638	17	US-09-351-224-2	Sequence 2, Appl
2	1638	100.0	1638	17	US-09-351-224E-2	Sequence 2, Appl
3	1638	100.0	1638	17	US-09-351-823-2	Sequence 2, Appl
4	1638	100.0	1638	26	US-09-677-488-2	Sequence 2, Appl
5	1638	100.0	1638	26	US-09-677-488A-2	Sequence 2, Appl
6	1638	100.0	1638	26	US-09-677-682-2	Sequence 2, Appl
7	1638	100.0	1638	26	US-09-677-682A-2	Sequence 2, Appl
8	1638	100.0	1638	26	US-09-677-682B-2	Sequence 2, Appl
9	1638	100.0	1638	33	US-09-882-694-2	Sequence 2, Appl
10	1638	100.0	1638	33	US-09-882-694A-2	Sequence 2, Appl
11	1575	96.2	1691	17	US-09-351-224-1	Sequence 1, Appl
12	1575	96.2	1691	17	US-09-351-823-1	Sequence 1, Appl
13	1575	96.2	1691	26	US-09-677-488-1	Sequence 1, Appl
14	1575	96.2	1691	26	US-09-677-488A-1	Sequence 1, Appl
15	1575	96.2	1691	26	US-09-677-682-1	Sequence 1, Appl
16	1575	96.2	1691	26	US-09-677-682A-1	Sequence 1, Appl
17	1575	96.2	1691	26	US-09-677-682B-1	Sequence 1, Appl
18	1575	96.2	1691	33	US-09-882-694-1	Sequence 1, Appl
19	1575	96.2	1691	33	US-09-882-694A-1	Sequence 1, Appl
20	1575	96.2	2660	52	US-60-138-103-3656	Sequence 3656, Ap
21	368.8	22.5	3246	18	US-09-404-520-981	Sequence 981, App
22	368.8	22.5	3246	18	US-09-404-520-20066	Sequence 20066, A
23	355.4	21.7	1672	18	US-09-404-520-3374	Sequence 3374, Ap
24	271.2	16.6	2362	18	US-09-404-520-3374	Sequence 3374, Ap
25	271.2	16.6	2362	18	US-09-404-520-3374	Sequence 3374, Ap
26	237.2	14.5	1981	52	US-60-138-103-4250	Sequence 4250, Ap
27	237.2	13.9	1590	19	US-09-504-358-21	Sequence 21, Appl
28	237.2	13.9	1590	35	US-09-504-358-21	Sequence 21, Appl
29	237.2	13.9	1590	35	US-09-504-358-21	Sequence 21, Appl
30	237.2	13.9	1590	51	US-60-120-702-21	Sequence 16, Appl
31	237.2	13.9	11471	19	US-09-504-358-16	Sequence 16, Appl
32	237.2	13.9	11471	35	US-09-504-358-16	Sequence 16, Appl

32 227.6 13.9 11475 51 US-60-120-702-16 Sequence 16, Appl  
33 215.6 13.2 1190 18 US-09-404-520-19705 Sequence 19705, A  
34 174 10.6 732 18 US-09-417-507-18452 Sequence 18452, A  
35 165.6 10.1 831 18 US-09-417-507-18449 Sequence 18449, A  
36 159.2 9.7 695 20 US-09-533-559-4967 Sequence 4967, Ap  
37 153.2 9.4 1644 25 US-09-648-004-9 Sequence 9, Appl  
38 153.2 9.4 17417 25 US-09-648-004-27 Sequence 27, Appl  
39 144.2 8.8 6879 18 US-09-404-520-5073 Sequence 5073, Ap  
40 140.6 8.6 1662 19 US-09-504-358-5 Sequence 5, Appl  
41 140.6 8.6 1662 35 US-09-954-314-5 Sequence 5, Appl  
42 140.6 8.6 1662 51 US-60-120-702-5 Sequence 5, Appl  
43 140.6 8.6 10629 19 US-09-504-358-15 Sequence 15, Appl  
44 140.6 8.6 10629 35 US-09-954-314-15 Sequence 15, Appl  
45 140.6 8.6 10631 51 US-60-120-702-15 Sequence 15, Appl

## ALIGNMENTS

RESULT 1  
US-09-351-224-2  
; Sequence 2, Application US/09351224  
; GENERAL INFORMATION:  
; APPLICANT: DuVick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351.224  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: flavin monooxygenase, coding sequence  
US-09-351-224-2

Query Match 100.0%; Score 1638; DB 17; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtcgccaccagcaactccagagcgattgttccgtcgcatgagcccatcgtt 60  
Db 1 atgtcgccaccagcaactccagagcgattgttccgtcgcatgagcccatcgtt 60  
Qy 61 ggagccgctcagcgcatctctgctgtgtacaaattgcgaagctcagactcaagcc 120  
Db 61 ggagccgctcagcgcatctctgctgtgtacaaattgcgaagctcagactcaagcc 120  
Qy 121 aaaaattcgaggagcccgagatttggcggtctggtgcaactggaacgctacccctggc 180  
Db 121 aaaaattcgaggagcccgagatttggcggtctggtgcaactggaacgctacccctggc 180  
Qy 181 gctcgtgtgattcgagagccctcttaccactgaacattcccggaagtatgaaagac 240  
Db 181 gctcgtgtgattcgagagccctcttaccactgaacattcccggaagtatgaaagac 240  
Qy 241 tggacctgtcttgcgtatcctgaccagaaagagtgtgctgcatatgttcaacctgt 300  
Db 241 tggacctgtcttgcgtatcctgaccagaaagagtgtgctgcatatgttcaacctgt 300  
Qy 301 gacaagatcgagggttgagaaagacgtctacttcggagctgaggtgttgatgcgcgg 360  
Db 301 gacaagatcgagggttgagaaagacgtctacttcggagctgaggtgttgatgcgcgg 360  
Qy 361 tatgccagagatctggcagctgactgtcaagacgtcggtggccatgttgcgacggca 420  
Db 361 tatgccagagatctggcagctgactgtcaagacgtcggtggccatgttgcgacggca 420  
Qy 421 aagtatctattctcgtacaggggttctccacaggaagacactcccgactcccccggc 480  
Db 421 aagtatctattctcgtacaggggttctccacaggaagacactcccgactcccccggc 480

Db 421 aagtatctattctcgtacaggggttctccacaggaagacactcccgactcccccggc 480  
Qy 481 ctgcgcgatttcaacggggaagtgatttcatttcgagtgctcgtggcaggaagacttcagcga 540  
Db 481 ctgcgcgatttcaacggggaagtgatttcatttcgagtgctcgtggcaggaagacttcagcga 540  
Qy 541 gagggccagagagtcgcgcgtcactcgtgctcggggccacagcatccagatttgcagag 600  
Db 541 gagggccagagagtcgcgcgtcactcgtgctcggggccacagcatccagatttgcagag 600  
Qy 601 ttggccaaagagctgcacagtaacctgtttatgcgaagcgagctatgtctgccc 660  
Db 601 ttggccaaagagctgcacagtaacctgtttatgcgaagcgagctatgtctgccc 660  
Qy 661 atgcygcaacgacgattgtaggaacgacacagacagcctggaagcctactaccaccag 720  
Db 661 atgcygcaacgacgattgtaggaacgacacagacagcctggaagcctactaccaccag 720  
Qy 721 ctgtttgaagcgagtcgaaagtctcgtgatttcccggtccagggcaccgctcgattggc 780  
Db 721 ctgtttgaagcgagtcgaaagtctcgtgatttcccggtccagggcaccgctcgattggc 780  
Qy 781 atctttgaagtcagcccgagcagcgaggcctatttcgaagagttgtggagcgtgagg 840  
Db 781 atctttgaagtcagcccgagcagcgaggcctatttcgaagagttgtggagcgtgagg 840  
Qy 841 gectttaatttcttgcgtccagtaaccgagagtcattgttgcacaaagggccacccga 900  
Db 841 gectttaatttcttgcgtccagtaaccgagagtcattgttgcacaaagggccacccga 900  
Qy 901 ctggtctatgacttctgggccaagaaactcgtatcgtcaatcccgccaagaga 960  
Db 901 ctggtctatgacttctgggccaagaaactcgtatcgtcaatcccgccaagaga 960  
Qy 961 gatctcattggtctccttgagccgcgtactggttctggtaccagcgctccaccatgag 1020  
Db 961 gatctcattggtctccttgagccgcgtactggttctggtaccagcgctccaccatgag 1020  
Qy 1021 agcgactactcgaattctggaacagccgagcgtcgaaattgtgaattcgaacaaatcg 1080  
Db 1021 agcgactactcgaattctggaacagccgagcgtcgaaattgtgaattcgaacaaatcg 1080  
Qy 1081 cccattgtggtcgttaccaaagacaggtgtgctcttggatgacggcaggaaggaatgc 1140  
Db 1081 cccattgtggtcgttaccaaagacaggtgtgctcttggatgacggcaggaaggaatgc 1140  
Qy 1141 gacacgactcgtcgtcgacaggggttctcgacagtttcaactggtcattgacacatatggc 1200  
Db 1141 gacacgactcgtcgtcgacaggggttctcgacagtttcaactggtcattgacacatatggc 1200  
Qy 1201 ttgaaaaaacacagcagcgagtcgaaactgaaagggtgtggaagaatggcattacttat 1260  
Db 1201 ttgaaaaaacacagcagcgagtcgaaactgaaagggtgtggaagaatggcattacttat 1260  
Qy 1261 atggggagttctctcatttggcttccccaatgctcttcttctgcccacggtcgaaccccg 1320  
Db 1261 atggggagttctctcatttggcttccccaatgctcttcttctgcccacggtcgaaccccg 1320  
Qy 1321 accgtctcttccaaagggcccaacgatcatagaaacccaagtcgaacttgatcgcgataca 1380  
Db 1321 accgtctcttccaaagggcccaacgatcatagaaacccaagtcgaacttgatcgcgataca 1380  
Qy 1381 attgcaaaagtgtgagggcgagcagccacgtccgttggggcgacgaaatccagcacaagag 1440  
Db 1381 attgcaaaagtgtgagggcgagcagccacgtccgttggggcgacgaaatccagcacaagag 1440  
Qy 1441 gcatggtcattatgattgcgaagtgaacagtcacactctgttcccttgacggttcg 1500  
Db 1441 gcatggtcattatgattgcgaagtgaacagtcacactctgttcccttgacggttcg 1500  
Qy 1501 tggtagactggaggcaacatccctgggaaagcaacacgtgcttttaaccttaagccggg 1560  
Db 1501 tggtagactggaggcaacatccctgggaaagcaacacgtgcttttaaccttaagccggg 1560

QY 1561 attgctctatgagcagatctgtcaagagaggtggccaatggatgggtttgatgtg 1620  
|||||  
Db 1561 attgctctatgagcagatctgtcaagagaggtggccaatgggtttgatgtg 1620  
  
QY 1621 attcatgctccctgctaa 1638  
|||||  
Db 1621 attcatgctccctgctaa 1638

## RESULT 2

US-09-351-224E-2

; Sequence 2, Application US/09351224E

; GENERAL INFORMATION:

; APPLICANT: Duwick, Jon

; APPLICANT: Maddox, Joyce

; APPLICANT: Gilliam, Jacob

; APPLICANT: Folkerts, Otto

; APPLICANT: Crasta, Oswald R.

; TITLE OF INVENTION: Compositions and Methods for Fumonisin

; FILE REFERENCE: Detoxification

; CURRENT APPLICATION NUMBER: US/09/351,224E

; CURRENT FILING DATE: 1999-07-12

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 1638

; TYPE: DNA

; ORGANISM: Exophiala spinifera

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (0)...(0)

; OTHER INFORMATION: flavin monooxygenase, fully spliced

US-09-351-224E-2

Query Match 100.0%; Score 1638; DB 17; Length 1638;  
Rest Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtggccacacagcaactccagagcgattgttcgcgcatgacgagccatcactgtt 60  
|||||  
Db 1 atgtggccacacagcaactccagagcgattgttcgcgcatgacgagccatcactgtt 60  
  
QY 61 ggagccggcctcagcgcatctctgctgtgtacaaattcgcaaatgcagactcaacgcc 120  
|||||  
Db 61 ggagccggcctcagcgcatctctgctgtgtacaaattcgcaaatgcagactcaacgcc 120  
  
QY 121 aaatcttcgaggagccccgattttggcggtctggtgcaactggaacccgctaccctggc 180  
|||||  
Db 121 aaatcttcgaggagccccgattttggcggtctggtgcaactggaacccgctaccctggc 180  
  
QY 181 gctggtgtgattcgagagccctctaccacatgaacattcccgagatggaagac 240  
|||||  
Db 181 gctggtgtgattcgagagccctctaccacatgaacattcccgagatggaagac 240  
  
QY 241 tggacctgtcttcgctatctgacacagagattgctgtcatatttccactgt 300  
|||||  
Db 241 tggacctgtcttcgctatctgacacagagattgctgtcatatttccactgt 300  
  
QY 301 gacaagatccggggctgagaaagacgtctacttcgagagtgaggttgatgcgagg 360  
|||||  
Db 301 gacaagatccggggctgagaaagacgtctacttcgagagtgaggttgatgcgagg 360  
  
QY 361 tatgcagagatctgggacactggacttcaagacgtcggtggccatgttgcgacggca 420  
|||||  
Db 361 tatgcagagatctgggacactggacttcaagacgtcggtggccatgttgcgacggca 420  
  
QY 421 aagatctcattctcgtacgggttgcctccacaggaagacacactccgcaactcccgcc 480  
|||||  
Db 421 aagatctcattctcgtacgggttgcctccacaggaagacacactccgcaactcccgcc 480



US-09-677-488-2

; Sequence 2, Application US/09677488

; GENERAL INFORMATION:

; APPLICANT: Duvick, Jon

; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification

; FILE REFERENCE: 35718/204100 (35718-111A)

; CURRENT APPLICATION NUMBER: US/09/677,488

; CURRENT FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: 09/351,224

; PRIOR FILING DATE: 1999-07-12

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1638

; TYPE: DNA

; ORGANISM: Exophiala spinifera

; FEATURE:

; OTHER INFORMATION: flavin monooxygenase, fully spliced cDNA

US-09-677-488-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcggccaccagcaactccagagcgatgttcctcgcatcgagcgccatcactgtt 60

Db 1 atgtcggccaccagcaactccagagcgatgttcctcgcatcgagcgccatcactgtt 60

QY 61 gtagccggcctcagcgcatctctgtgtacaaattgcgaatgcgaagctcagactcaagcc 120

Db 61 gtagccggcctcagcgcatctctgtgtacaaattgcgaatgcgaagctcagactcaagcc 120

QY 121 aaaaattcctcaggagccccagattttggcgctgtctggcactggaacccctaccctggc 180

Db 121 aaaaattcctcaggagccccagattttggcgctgtctggcactggaacccctaccctggc 180

QY 181 gctcgtgtgattcggagagccctctaccactgaacattcccgaggtatggaagac 240

Db 181 gctcgtgtgattcggagagccctctaccactgaacattcccgaggtatggaagac 240

QY 241 tggactgttcttcgctatctcaccagaaagattgctgtcatatgttcaccactgt 300

Db 241 tggactgttcttcgctatctcaccagaaagattgctgtcatatgttcaccactgt 300

QY 301 gacaagatccggggtctgagaaagacgtctacttcggagctgaggtgttgatgcgag 360

Db 301 gacaagatccggggtctgagaaagacgtctacttcggagctgaggtgttgatgcgag 360

QY 361 tatgcagagatctgggacactggactgtcaagacgtcggtggccatgtttgcgagcca 420

Db 361 tatgcagagatctgggacactggactgtcaagacgtcggtggccatgtttgcgagcca 420

QY 421 aagatctcattcgtcaggggtgtctccacaggaagcacctcccgcaactcccgcc 480

Db 421 aagatctcattcgtcaggggtgtctccacaggaagcacctcccgcaactcccgcc 480

QY 481 ctgcgcatttcaacggaggggtgattcatttcgagtcctggcgacgaagacttcgagcca 540

Db 481 ctgcgcatttcaacggaggggtgattcatttcgagtcctggcgacgaagacttcgagcca 540

QY 541 gagggccagagatcgccgcatcgtgtccggggccacaagcatccagattgttcaggag 600

Db 541 gagggccagagatcgccgcatcgtgtccggggccacaagcatccagattgttcaggag 600

QY 601 ttggccagagaggtcaccaggttaaccatgtttatcgaaagccgagctattgtctccc 660

Db 601 ttggccagagaggtcaccaggttaaccatgtttatcgaaagccgagctattgtctccc 660

QY 661 atcggccacgaacagatggataggaacgaacagacagcctggaagccctactaccacag 720

Db 661 atcggccacgaacagatggataggaacgaacagacagcctggaagccctactaccacag 720

QY 721 ctgtttgaagcgagtcgaaagtctcgattcccggtcccgagccaccctcggttggc 780

Db 721 ctgtttgaagcgagtcgaaagtctcgattcccggtcccgagccaccctcggttggc 780

QY 781 atctttgaagtcagcccgagcagcgagggagcctatttcgaagagttgtggagcgtggg 840

Db 781 atctttgaagtcagcccgagcagcgagggagcctatttcgaagagttgtggagcgtggg 840

QY 841 gctttaattttctgtgttcgagtaaccgagagatcatgtttgacaaaaagggccaacgga 900

Db 841 gctttaattttctgtgttcgagtaaccgagagatcatgtttgacaaaaagggccaacgga 900

QY 901 ctggcttatgacttctggggccaaaagactcgtatcgtcatcgtcaatcccggaagaga 960

Db 901 ctggcttatgacttctggggccaaaagactcgtatcgtcatcgtcaatcccggaagaga 960

QY 961 gatctcatggtcctctgagcccgctactggttcggtaccgaagcgtccccactggag 1020

Db 961 gatctcatggtcctctgagcccgctactggttcggtaccgaagcgtccccactggag 1020

QY 1021 agcgactactacgaaatctggacaaagccgagcgtcgaaattgtgaattctagaacaatcg 1080

Db 1021 agcgactactacgaaatctggacaaagccgagcgtcgaaattgtgaattctagaacaatcg 1080

QY 1081 cccatttggctgtttacaaaagacaggtgtctctttagtgagcgagcgaaggggaatgac 1140

Db 1081 cccatttggctgtttacaaaagacaggtgtctctttagtgagcgagcgaaggggaatgac 1140

QY 1141 gacacgactcgtctgagcgagcgttttcgacagtttcacttggctcatttgacacatatggc 1200

Db 1141 gacacgactcgtctgagcgagcgttttcgacagtttcacttggctcatttgacacatatggc 1200

QY 1201 ttgaaaaacaaacagagagtggaactgaagaggtgtgaaagatggacatatctacttat 1260

Db 1201 ttgaaaaacaaacagagagtggaactgaagaggtgtgaaagatggacatatctacttat 1260

QY 1261 atgggagttctctctatggtttccccaatgctcttctctcgtccacggctcaagcccg 1320

Db 1261 atgggagttctctctatggtttccccaatgctcttctctcgtccacggctcaagcccg 1320

QY 1321 accgtctcttccaaagcccaacgatcatagaaacccaagctgacttgatcccgataca 1380

Db 1321 accgtctcttccaaagcccaacgatcatagaaacccaagctgacttgatcccgataca 1380

QY 1381 attgcaaatgttgagccgagcagccacgtccgtgtgagcgagcgaatcagcaacagag 1440

Db 1381 attgcaaatgttgagccgagcagccacgtccgtgtgagcgagcgaatcagcaacagag 1440

QY 1441 gcattgctgattatgattgcgaagatgaacgagacactctgttcccttgacggattcg 1500

Db 1441 gcattgctgattatgattgcgaagatgaacgagacactctgttcccttgacggattcg 1500

QY 1501 tggtagactggaggcaacatccctgggaaagcaaacacgtctttaaccttcatagcggg 1560

Db 1501 tggtagactggaggcaacatccctgggaaagcaaacacgtctttaaccttcatagcggg 1560

QY 1561 attgctctctatgagcaatctctcaagagagaggtggccaatggatgggtttgatgtg 1620

Db 1561 attgctctctatgagcaatctctcaagagagaggtggccaatggatgggtttgatgtg 1620

QY 1621 ctctcgtcctcctgctaa 1638

Db 1621 ctctcgtcctcctgctaa 1638

RESULT 5

US-09-677-488A-2

; Sequence 2, Application US/09677488A

; GENERAL INFORMATION:

; APPLICANT: Duvick, Jon

; APPLICANT: Maddox, Joyce

; APPLICANT: Gilliam, Jacob

; APPLICANT: Folkerts, Otto

APPLICANT: Crasta, Oswald R.  
TITLE OF INVENTION: Compositions and Methods for Fumonisin  
FILE REFERENCE: 35718/204100  
CURRENT APPLICATION NUMBER: US/09/677,488A  
CURRENT FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 09/351,224  
PRIOR FILING DATE: 1999-07-12  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1638  
TYPE: DNA  
ORGANISM: Exophiala spinifera  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)..(0)  
OTHER INFORMATION: Flavin monooxygenase, fully spliced

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtggccacagcaactcagagggcgtatgttcctgcgatgcgacgcgcacatcgtt 60  
Db 1 atgtggccacagcaactcagagggcgtatgttcctgcgatgcgacgcgcacatcgtt 60  
QY 61 gaggccggtcagcgacatctcgtgtgtacaaattgcgaagtcgaagtcagactcaagcc 120  
Db 61 gaggccggtcagcgacatctcgtgtgtacaaattgcgaagtcgaagtcagactcaagcc 120  
QY 121 aaaaatttcgagggagccccgatttggcgcgctgtggactggaaacgcgtacccctggc 180  
Db 121 aaaaatttcgagggagccccgatttggcgcgctgtggactggaaacgcgtacccctggc 180  
QY 181 gctcgtgtgtatcggagagccctctaccactgaacattcccgaaagtcgaaagac 240  
Db 181 gctcgtgtgtatcggagagccctctaccactgaacattcccgaaagtcgaaagac 240  
QY 241 tggacctgttgcctcactcctgacccagaaagtgctgtcatatgttccaccactgt 300  
Db 241 tggacctgttgcctcactcctgacccagaaagtgctgtcatatgttccaccactgt 300  
QY 301 gacaagatccgggcttgaaagagcgtctacttcgaagctgaggtgtgtatgcgcgg 360  
Db 301 gacaagatccgggcttgaaagagcgtctacttcgaagctgaggtgtgtatgcgcgg 360  
QY 361 tatgccagagatctgggacactggactgtcaagacgtcggtgtggccatgttgcgaagcga 420  
Db 361 tatgccagagatctgggacactggactgtcaagacgtcggtgtggccatgttgcgaagcga 420  
QY 421 aagtatctcattcgtacaggggttgcctccacagaaagacactcccgacactcccgcc 480  
Db 421 aagtatctcattcgtacaggggttgcctccacagaaagacactcccgacactcccgcc 480  
QY 481 ctgcgcatcttcacaggaaggtgattcattcgtgctggcagcaagacttcgacgca 540  
Db 481 ctgcgcatcttcacaggaaggtgattcattcgtgctggcagcaagacttcgacgca 540  
QY 541 gaggccagagatcgccgtcactcgtgcggggccacaaagcatccagattgttcagag 600  
Db 541 gaggccagagatcgccgtcactcgtgcggggccacaaagcatccagattgttcagag 600  
QY 601 ttggccaaagaggtcagcaggttaaccattgttatgcgaagccgagctattgtctgcc 660  
Db 601 ttggccaaagaggtcagcaggttaaccattgttatgcgaagccgagctattgtctgcc 660  
QY 661 atgcggcaacagacgatggataggaacagacagacagcgtggaagggcctactaccacacg 720  
Db 661 atgcggcaacagacgatggataggaacagacagacagcgtggaagggcctactaccacacg 720

## RESULT 6

US-09-677-682-2

; Sequence 2, Application US/09677682

; GENERAL INFORMATION:

; APPLICANT: Duvick, Jon

; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification

; FILE REFERENCE: 35718/204101 (5718-111b)

; CURRENT APPLICATION NUMBER: US/09/677,682

QY 721 ctgttgagcagcagtcgaaagtctcgattcgattcccggtccagcaccctcgttggc 780  
Db 721 ctgttgagcagcagtcgaaagtctcgattcgattcccggtccagcaccctcgttggc 780  
QY 781 atctttgaagtacgccccgagcagcgaggcgctatttcgaagagttgtggagcgctgg 840  
Db 781 atctttgaagtacgccccgagcagcgaggcgctatttcgaagagttgtggagcgctgg 840  
QY 841 gcttttaattttctgtcctcagtagcaggaagtcagtggttgacaaaagggccaaacga 900  
Db 841 gcttttaattttctgtcctcagtagcaggaagtcagtggttgacaaaagggccaaacga 900  
QY 901 ctggtctatgactctcggccaaaagactcgtctcgtatcgtcaatcccgcaagaga 960  
Db 901 ctggtctatgactctcggccaaaagactcgtctcgtatcgtcaatcccgcaagaga 960  
QY 961 gatctcatggctcctctggagcccgctactggttcgggtaccagcgctcccccctggag 1020  
Db 961 gatctcatggctcctctggagcccgctactggttcgggtaccagcgctcccccctggag 1020  
QY 1021 agcactactacgaatgctggacaagccgagcgtcgaatctgtaattctagaacaatcg 1080  
Db 1021 agcactactacgaatgctggacaagccgagcgtcgaatctgtaattctagaacaatcg 1080  
QY 1081 cccattgtgctgttcaaaagacaggtgtgctcttgtagtgacggcgacagagggaaatgc 1140  
Db 1081 cccattgtgctgttcaaaagacaggtgtgctcttgtagtgacggcgacagagggaaatgc 1140  
QY 1141 gacacgctcgtgctggcagcgggttccgacagtttccactggtcctatgacacatatgggc 1200  
Db 1141 gacacgctcgtgctggcagcgggttccgacagtttccactggtcctatgacacatatgggc 1200  
QY 1201 ttgaaaaaagcagcagggagtgacctgaagaggtgtggaagagtgacatatctactat 1260  
Db 1201 ttgaaaaaagcagcagggagtgacctgaagaggtgtggaagagtgacatatctactat 1260  
QY 1261 atggagctctctcctcatggcttcccccaatgctctcttcgtccgcagcgtccaaagcccg 1320  
Db 1261 atggagctctctcctcatggcttcccccaatgctctcttcgtccgcagcgtccaaagcccg 1320  
QY 1321 accgtcttcccaacggcccaacgacatcagaaacccaagtcgacttgatcgccgacata 1380  
Db 1321 accgtcttcccaacggcccaacgacatcagaaacccaagtcgacttgatcgccgacata 1380  
QY 1381 attcaagcttgagcggccagcagccagctccgttgagggcagcgaatcagcacaagag 1440  
Db 1381 attcaagcttgagcggccagcagccagctccgttgagggcagcgaatcagcacaagag 1440  
QY 1441 gcatgctcgtattatgattcccaagatgaacgagcagacactctgttcccttgacggattcg 1500  
Db 1441 gcatgctcgtattatgattcccaagatgaacgagcagacactctgttcccttgacggattcg 1500  
QY 1501 tgggtgactggagggcaacatccctgggaagcacaacgctgctttaacctctataggcggg 1560  
Db 1501 tgggtgactggagggcaacatccctgggaagcacaacgctgctttaacctctataggcggg 1560  
QY 1561 attgctctctatgagcagatctgttcaagagaaggtggccaattgggatttgatgtg 1620  
Db 1561 attgctctctatgagcagatctgttcaagagaaggtggccaattgggatttgatgtg 1620  
QY 1621 ctctcatcctcctgctaaa 1638  
Db 1621 ctctcatgctcctcgtaaa 1638

RESULT 6

US-09-677-682-2

; Sequence 2, Application US/09677682

; GENERAL INFORMATION:

; APPLICANT: Duvick, Jon

; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification

; FILE REFERENCE: 35718/204101 (5718-111b)

; CURRENT APPLICATION NUMBER: US/09/677,682

; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: flavin monooxygenase, fully spliced cDNA  
US-09-677-682-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtcgccacgaacactccagagcgatgttcctgcgcacatcgagcccatcgtt 60  
Db 1 atgtcgccacgaacactccagagcgatgttcctgcgcacatcgagcccatcgtt 60

Qy 61 ggaagccgctcagggcactctctgtgtgtacaaattgcgaagctcagactcaacgcc 120  
Db 61 ggaagccgctcagggcactctctgtgtgtacaaattgcgaagctcagactcaacgcc 120

Qy 121 aaaaattctcagggagcccgatgttgccggcgctctggcactggaacccgtaccctggc 180  
Db 121 aaaaattctcagggagcccgatgttgccggcgctctggcactggaacccgtaccctggc 180

Qy 181 gctcgtgtgattcggagagccctcttaccacactgaacattcccgagtggaagac 240  
Db 181 gctcgtgtgattcggagagccctcttaccacactgaacattcccgagtggaagac 240

Qy 241 tggacgtgtcttgcgcgtatctcgtacagaaagattgctgtcatattgttccaccactgt 300  
Db 241 tggacgtgtcttgcgcgtatctcgtacagaaagattgctgtcatattgttccaccactgt 300

Qy 301 gacaagatccgggcttgagaaacactcttaccagagctgaggtgagtgatgcgcgg 360  
Db 301 gacaagatccgggcttgagaaacactcttaccagagctgaggtgagtgatgcgcgg 360

Qy 361 tatgcagagatctgggacactggactgtcaagacgtcgctggccatgttgcgacggca 420  
Db 361 tatgcagagatctgggacactggactgtcaagacgtcgctggccatgttgcgacggca 420

Qy 421 aagtatctcattctcgtcaggggtgtctccacagaaagacactcccgcaactcccggc 480  
Db 421 aagtatctcattctcgtcaggggtgtctccacagaaagacactcccgcaactcccggc 480

Qy 481 ctgcgcgatttaacagggaagtgatttcatttcgagtgctggcagaaagacttcgacgca 540  
Db 481 ctgcgcgatttaacagggaagtgatttcatttcgagtgctggcagaaagacttcgacgca 540

Qy 541 gagggccagagatcgccgtcactcgtgtcgggggccaagcatcagattgttcaggag 600  
Db 541 gagggccagagatcgccgtcactcgtgtcgggggccaagcatcagattgttcaggag 600

Qy 601 ttggccaagaagctaacaggttaaccatgtttatgcgaagccgagctattgtctccc 660  
Db 601 ttggccaagaagctaacaggttaaccatgtttatgcgaagccgagctattgtctccc 660

Qy 661 atcgccgaacgaacgatgataaggaacgaacagacagctggaagggcctactaccacag 720  
Db 661 atcgccgaacgaacgatgataaggaacgaacagacagctggaagggcctactaccacag 720

Qy 721 ctgtttgaacgaagctcgaagctcgaattgttcctccggtccagaccgtcgttggc 780  
Db 721 ctgtttgaacgaagctcgaagctcgaattgttcctccggtccagaccgtcgttggc 780

Qy 781 atctttgaagtcagccccagcagcgggagggcctatttcgaagagttgtggagcgtggg 840  
Db 781 atctttgaagtcagccccagcagcgggagggcctatttcgaagagttgtggagcgtggg 840

Qy 841 gcctttaattttcttgcctaccgagagtcattggttgacaaaaagcgcaaccga 900  
Db 841 gcctttaattttcttgcctaccgagagtcattggttgacaaaaagcgcaaccga 900

Qy 901 ctggtctatgacttctgggcccacaaagactcgtatctcgtatcgtcgaatcggaagaga 960  
Db 901 ctggtctatgacttctgggcccacaaagactcgtatctcgtatcgtcgaatcggaagaga 960

Qy 961 gatctcattgctcctcgtggagcccgctactggttcggtaccagcgtcccccactggag 1020  
Db 961 gatctcattgctcctcgtggagcccgctactggttcggtaccagcgtcccccactggag 1020

Qy 1021 agcgactactcgaatactgtgacaaagccgagcgctgaaattgtgaattcgaatacgaacacg 1080  
Db 1021 agcgactactcgaatactgtgacaaagccgagcgctgaaattgtgaattcgaatacgaacacg 1080

Qy 1081 cccatttggtctgttacaagaacagagtgctctcttggagtgacggcagcaagaggggaatgc 1140  
Db 1081 cccatttggtctgttacaagaacagagtgctctcttggagtgacggcagcaagaggggaatgc 1140

Qy 1141 gacacgactcgtcgtcggcgagcggttctcagacagtttccactggtcattgcacatatgggc 1200  
Db 1141 gacacgactcgtcgtcggcgagcggttctcagacagtttccactggtcattgcacatatgggc 1200

Qy 1201 ttgaaaaacaaagcagcgagtggaactgaagagagtggtggaagatggcatctacttat 1260  
Db 1201 ttgaaaaacaaagcagcgagtggaactgaagagagtggtggaagatggcatctacttat 1260

Qy 1261 atggagactctctcattcgtctcccaatgctcttctcgtccacggctcgaagcccg 1320  
Db 1261 atggagactctctcattcgtctcccaatgctcttctcgtccacggctcgaagcccg 1320

Qy 1321 accgtctcttccaaagcccaacgatcatagaacccaagtcgaactgtatgcgcgataca 1380  
Db 1321 accgtctcttccaaagcccaacgatcatagaacccaagtcgaactgtatgcgcgataca 1380

Qy 1381 attgcaaaagtggagcgagcagccacgctccgttgaggcgacgaaatcagcacaaag 1440  
Db 1381 attgcaaaagtggagcgagcagccacgctccgttgaggcgacgaaatcagcacaaag 1440

Qy 1441 gcatggtcattatgattgcgaagatgaacgagacacactctgttcccttgcagggattcg 1500  
Db 1441 gcatggtcattatgattgcgaagatgaacgagacacactctgttcccttgcagggattcg 1500

Qy 1501 tggtagactggaggcaacatccctgggaaagcaacacgctctttaaacttcattagcggg 1560  
Db 1501 tggtagactggaggcaacatccctgggaaagcaacacgctctttaaacttcattagcggg 1560

Qy 1561 attgctctctatgacagatctgtcaagagaagtgcccaattgggattgttgatgtg 1620  
Db 1561 attgctctctatgacagatctgtcaagagaagtgcccaattgggattgttgatgtg 1620

Qy 1621 ctctcatgctccctgctaa 1638  
Db 1621 ctctcatgctccctgctaa 1638

RESULT 7  
US-09-677-682A-2  
; Sequence 2, Application US/09677682A  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; TITLE OF INVENTION: Detoxification  
; FILE REFERENCE: 35718/204101  
; CURRENT APPLICATION NUMBER: US/09/677,682A  
; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224

; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Flavin monooxygenase, fully spliced  
US-09-677-682A-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtcggccaccagcaactccagagcgattgttcctcgatcgacgagccatcatcgtt 60  
Db 1 atgtcggccaccagcaactccagagcgattgttcctcgatcgacgagccatcatcgtt 60

Qy 61 ggaagccggctcagcgcatctctgtgtacaaattgcgaagctcagactcaagcgc 120  
Db 61 ggaagccggctcagcgcatctctgtgtacaaattgcgaagctcagactcaagcgc 120

Qy 121 aaatcttcgagagagccccaatttggcggtctgacgtgacgtgaacccgtaccctggc 180  
Db 121 aaatcttcgagagagccccaatttggcggtctgacgtgacgtgaacccgtaccctggc 180

Qy 181 gctcgtgtgattcgagagccctctaccactgaacattcccgaaagtatggaagac 240  
Db 181 gctcgtgtgattcgagagccctctaccactgaacattcccgaaagtatggaagac 240

Qy 241 tggacctcgtctgcgctatctcgtacacagaagagttgtgtgtatgttcacactgt 300  
Db 241 tggacctcgtctgcgctatctcgtacacagaagagttgtgtgtatgttcacactgt 300

Qy 301 gacaagatccgggcttgagaaagcgtctacttcgagctgaggtgtgtgtgctgcgg 360  
Db 301 gacaagatccgggcttgagaaagcgtctacttcgagctgaggtgtgtgtgctgcgg 360

Qy 361 tatccagagatctgggacactggaactgcaagacgtcggtggcactgttcgacagcga 420  
Db 361 tatccagagatctgggacactggaactgcaagacgtcggtggcactgttcgacagcga 420

Qy 421 aagtatctacttcgactcgggtgtgtccacaggaagcacactcccgactcccgccg 480  
Db 421 aagtatctacttcgactcgggtgtgtccacaggaagcacactcccgactcccgccg 480

Qy 481 ctgcgcgatttcaacgggaaggtgatttcattcgagtgctggcagcaagacttcgacgca 540  
Db 481 ctgcgcgatttcaacgggaaggtgatttcattcgagtgctggcagcaagacttcgacgca 540

Qy 541 gaggggccagagatcgccgtcactcgggtgcgggcccacaaagcatccagattgttcagag 600  
Db 541 gaggggccagagatcgccgtcactcgggtgcgggcccacaaagcatccagattgttcagag 600

Qy 601 ttggccaagaagcgtgaccaggttaaccatgtttatcgaaagccgagctattgtctgcc 660  
Db 601 ttggccaagaagcgtgaccaggttaaccatgtttatcgaaagccgagctattgtctgcc 660

Qy 661 atgcggcaacgacgattgtaggaacacagacagcgtggaagcctactaccacacg 720  
Db 661 atgcggcaacgacgattgtaggaacacagacagcgtggaagcctactaccacacg 720

Qy 721 ctgtttgaagcagtgtaagctcggattgagttcccggtccagagccctcggtggc 780  
Db 721 ctgtttgaagcagtgtaagctcggattgagttcccggtccagagccctcggtggc 780

Qy 781 atctttgaagtcagccccagcagcgggagcgtctatttcgaagagttgtgagcgtggg 840  
Db 781 atctttgaagtcagccccagcagcgggagcgtctatttcgaagagttgtgagcgtggg 840

Qy 841 gcctttaattttcttctgtccagttaccgagagatcattgttgacaaaaggccaaaccga 900  
Db 841 gcctttaattttcttctgtccagttaccgagagatcattgttgacaaaaggccaaaccga 900

Qy 901 ctggtctatgactctctggggccaaaagactcgatctctatctcgtcaatcccgccaaagaga 960  
Db 901 ctggtctatgactctctggggccaaaagactcgatctctatctcgtcaatcccgccaaagaga 960

Qy 961 gatctcatgctcctctctggagcgcgtactgggttcgtgtacaaagcgtctccctctggag 1020  
Db 961 gatctcatgctcctctctggagcgcgtactgggttcgtgtacaaagcgtctccctctggag 1020

Qy 1021 agcgactactacgaaatgctggacagccgagcgtcgaaattgtgaatctagaacaatcg 1080  
Db 1021 agcgactactacgaaatgctggacagccgagcgtcgaaattgtgaatctagaacaatcg 1080

Qy 1081 cccattgtgctctgttataaagacaggtgtgtctcttgaatgacgagcagcagaggaatgc 1140  
Db 1081 cccattgtgctctgttataaagacaggtgtgtctcttgaatgacgagcagcagaggaatgc 1140

Qy 1141 gacacgactcgtctgggacgggttttcgacagtttctactggtctcattgacacatatgggc 1200  
Db 1141 gacacgactcgtctgggacgggttttcgacagtttctactggtctcattgacacatatgggc 1200

Qy 1201 ttgaaaaaacacagcagcggagtgacactgaagaggtgtgtgaaagatggcattacttat 1260  
Db 1201 ttgaaaaaacacagcagcggagtgacactgaagaggtgtgtgaaagatggcattacttat 1260

Qy 1261 atgggagcttctctcatggtctccccaatgcctctctctcgtccacagcgtccaaagcccg 1320  
Db 1261 atgggagcttctctcatggtctccccaatgcctctctctcgtccacagcgtccaaagcccg 1320

Qy 1321 accgctcttccaaagcggcccaacatcatgaaacccaagtcgacttgatcgccgataca 1380  
Db 1321 accgctcttccaaagcggcccaacatcatgaaacccaagtcgacttgatcgccgataca 1380

Qy 1381 attgcaaaagtggaggccgagcagccacgtccgtttgaggcgacgaatacagcaagag 1440  
Db 1381 attgcaaaagtggaggccgagcagccacgtccgtttgaggcgacgaatacagcaagag 1440

Qy 1441 gcatggtcgtattatgattgccaagatgaacgagcacactctgttccctctgacggatcg 1500  
Db 1441 gcatggtcgtattatgattgccaagatgaacgagcacactctgttccctctgacggatcg 1500

Qy 1501 tgggtgactggaggaacacatccctgggaaagcaacacagtgctttaacacttcataaggcgg 1560  
Db 1501 tgggtgactggaggaacacatccctgggaaagcaacacagtgctttaacacttcataaggcgg 1560

Qy 1561 attgctctctatgacgagatctgttcaagagaggtggccaattgggattgggtttgatgtg 1620  
Db 1561 attgctctctatgacgagatctgttcaagagaggtggccaattgggattgggtttgatgtg 1620

Qy 1621 ctctcatgctccctgctaa 1638  
Db 1621 ctctcatgctccctgctaa 1638

## RESULT 8

US-09-677-682B-2  
; Sequence 2, Application us/09677682B  
; GENERAL INFORMATION:  
; APPLICANT: DuVick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 35718/204101  
; CURRENT APPLICATION NUMBER: US/09/677.682B  
; PRIOR APPLICATION NUMBER: 09/351,224



```
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: flavin monooxygenase, fully spliced
US-09-677-682B-2

Query Match      100.0%; Score 1638; DB 26; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgctggccaccagcaactccagagagatgttccgtcgcatgagcgccatcatcggt 60
DB 1 atgctggccaccagcaactccagagagatgttccgtcgcatgagcgccatcatcggt 60
QY 61 gggcgcgctcagcggaatctctgtgtgtacaaattggaaagctcagactcaacgcc 120
DB 61 gggcgcgctcagcggaatctctgtgtgtacaaattggaaagctcagactcaacgcc 120
QY 121 azaatcttcgagggagcccgatgttggcggtcttgcaactggaaacgctacctggc 180
DB 121 azaatcttcgagggagcccgatgttggcggtcttgcaactggaaacgctacctggc 180
QY 181 gctcgtgtgattcggagagcgccctctaccacactgaacatcccgaaagtatgaaagac 240
DB 181 gctcgtgtgattcggagagcgccctctaccacactgaacatcccgaaagtatgaaagac 240
QY 241 tggacgtgtcttgcgcgtatctgacacagaaagtgtgctgcatatgttccaccactgt 300
DB 241 tggacgtgtcttgcgcgtatctgacacagaaagtgtgctgcatatgttccaccactgt 300
QY 301 gacagatccgggcttgagaaagacgtctacttcgagcgtgaggtggttgatgcgcg 360
DB 301 gacagatccgggcttgagaaagacgtctacttcgagcgtgaggtggttgatgcgcg 360
QY 361 tatgccagagatctgggacactggactgtcaagacgtggtcgccatgttgcgacggca 420
DB 361 tatgccagagatctgggacactggactgtcaagacgtggtcgccatgttgcgacggca 420
QY 421 aagtatctctattctgcgtacgggtgtgtccacaggaagacactcccgcaactcccgcc 480
DB 421 aagtatctctattctgcgtacgggtgtgtccacaggaagacactcccgcaactcccgcc 480
QY 481 ctgcgcgatttcaacgggaagtgtatttcgagtcgctgtgcaagaaacttcgacga 540
DB 481 ctgcgcgatttcaacgggaagtgtatttcgagtcgctgtgcaagaaacttcgacga 540
QY 541 gaggccagagagtcgctgcgtatcgttccggggcccaacagcatccagattgttcaggag 600
DB 541 gaggccagagagtcgctgcgtatcgttccggggcccaacagcatccagattgttcaggag 600
QY 601 ttggcaagaagctgacagatgaacacatgtttatgcaaggccgagctattgtctgcc 660
DB 601 ttggcaagaagctgacagatgaacacatgtttatgcaaggccgagctattgtctgcc 660
QY 661 atcggcacaacgaatgataggaacgaacagacagcctggaaggcctactaccacacg 720
DB 661 atcggcacaacgaatgataggaacgaacagacagcctggaaggcctactaccacacg 720
QY 721 ctgtttgaagcagagtcgaagtcgtgatttgattcccggtccagggcacctcggttggc 780
DB 721 ctgtttgaagcagagtcgaagtcgtgatttgattcccggtccagggcacctcggttggc 780
QY 781 atctttgaagtcagcccgagcagcgaggccctatttcgaagattgtggagcgtggg 840
DB 781 atctttgaagtcagcccgagcagcgaggccctatttcgaagattgtggagcgtggg 840
```

## RESULT 9

```
US-09-882-694-2
; Sequence 2, Application US/09882694
; GENERAL INFORMATION:
; APPLICANT: Duwick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 35718/208255
; CURRENT APPLICATION NUMBER: US/09/882,694
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 09/351,224
```

```

; PRIOR FILING DATE: 1999-07-12
;
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1638
;   TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: flavin monooxygenase,
US-09-882-694-2

```

Query Match		100.0%	Score 1638;	DB 33;	Length 1638;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1638;		Conservative	0;	Mismatches	0;
QY	1	atgtcggccacacgaactccagagcgatgttcctgcgcgatcgacgcgcatacgtt	60		
DB	1	atgtcggccacacgaactccagagcgatgttcctgcgcgatcgacgcgcatacgtt	60		
QY	61	ggagcggcctcagcgcatctcgtgtacaaatgcgaagctcaagactcaagcgc	120		
DB	61	ggagcggcctcagcgcatctcgtgtacaaatgcgaagctcaagactcaagcgc	120		
QY	121	aaaattctcgaggagccccgaattttggcggcgctgtggcaactggaaacgcctaccctggc	180		
DB	121	aaaattctcgaggagccccgaattttggcggcgctgtggcaactggaaacgcctaccctggc	180		
QY	181	gctcgtgttgattcggagacgccttctaccaactgaacattcccgaaagtatggaaagac	240		
DB	181	gctcgtgttgattcggagacgccttctaccaactgaacattcccgaaagtatggaaagac	240		
QY	241	tggacctgttctccgctactctgaccgaagaagtttgtctcatatgtttccaccctgt	300		
DB	241	tggacctgttctccgctactctgaccgaagaagtttgtctcatatgtttccaccctgt	300		
QY	301	gacaagatccggggcttgaaaaagacgtctacttcggagctgaggtgggtgtatgcgcgg	360		
DB	301	gacaagatccggggcttgaaaaagacgtctacttcggagctgaggtgggtgtatgcgcgg	360		
QY	361	tatgcacagatctgggcaactggactctcaagacgtcggctggccattgttgcagcgca	420		
DB	361	tatgcacagatctgggcaactggactctcaagacgtcggctggccattgttgcagcgca	420		
QY	421	aagtatctcatctcgtacgggttgtctccacgaaagacacactcccgcactcccgcgg	480		
DB	421	aagtatctcatctcgtacgggttgtctccacgaaagacacactcccgcactcccgcgg	480		
QY	481	ctcgcgattccacggaagtgattcatttcgagtgctggacggaagacttcagcgca	540		
DB	481	ctcgcgattccacggaagtgattcatttcgagtgctggacggaagacttcagcgca	540		
QY	541	gagggccagagatgcgcgttcactcgtatgcggggccacaagcatccagattgttccaggag	600		
DB	541	gagggccagagatgcgcgttcactcgtatgcggggccacaagcatccagattgttccaggag	600		
QY	601	tggcccaagaaggtgaccaggttaaccattgtttatgcgaagccgagctattgtctgcgc	660		
DB	601	tggcccaagaaggtgaccaggttaaccattgtttatgcgaagccgagctattgtctgcgc	660		
QY	661	atgcgcaacgaatcgattaggaacgaacagacagctgtgaagcgctactaccaccag	720		
DB	661	atgcgcaacgaatcgattaggaacgaacagacagctgtgaagcgctactaccaccag	720		
QY	721	ctgtttgaagcgatcgaaaagtctcgatttgatttcccggtccagggcacccgtctggctggc	780		
DB	721	ctgtttgaagcgatcgaaaagtctcgatttgatttcccggtccagggcacccgtctggctggc	780		
QY	781	atctttgaagtcagccccagcagcgaggagcgcctatttcgaagagtttgtggagacgttggg	840		
DB	781	atctttgaagtcagccccagcagcgaggagcgcctatttcgaagagtttgtggagacgttggg	840		

Qy	841	g c t t t a a t t t t c t t g c c a g t c c c g a g a a g t c a t g g t t g a c a a a a a g c c a a c c g a	900
Dd	841	g c t t t a a t t t t c t t g c c a g t c c c g a g a a g t c a t g g t t g a c a a a a g c c a a c c g a	900
Qy	901	c t g g t c a t g a c t t c t g g g c c a a a a g a b c t g a t c t g t a t c g t c a a t c c g g c a a a g a g a	960
Dd	901	c t g g t c a t g a c t t c t g g g c c a a a a a g a c t g a t c t g t a t c g t c a a t c c g g c a a a g a g a	960
Qy	961	g a t c a t g g c t c t c t g a a c c g c g t a c t g g t t c g a t a c c a a c g c t c c c a c t g a g	1020
Dd	961	g a t c a t g g c t c t c t g a a c c g c g t a c t g g t t c g g t t c g g t a c c a a a g c g t c c c a c t g g a g	1020
Qy	1021	a g c a c t a c t a c a a a t g c t g s a a s c c g a g c g t g e g a a t t g t g a a t c t a g a a c a a t c g	1080
Dd	1021	a g c a c t a c t a c a a a t g c t g g a c a a c c g a g c g t c g a a a t t g t a a t c t a g a a c a a t c g	1080
Qy	1081	c c e a t t g g c t g t a c a a a g a c a g t g t g c t t t g a t g a c g g c a g a a g a g g a a t c g	1140
Dd	1081	c c e a t t g g c t g t a c a a a g a c a g t g t g c t t t g a t g a c g g c a g a a g a g g a a t c g	1140
Qy	1141	g a c a c a t c g t c t g g c a c g g t t c g a c a g t t c a c t g g c t c a t t g a c a c a t a t g g c	1200
Dd	1141	g a c a c a t c g t c t g g c a c g g t t c g a c a a g t t c a c t g g c t c a t t g a c a c a t a t g g c	1200
Qy	1201	t t g a a a a a a a g c a c g a g t g g a c c t g a a g g g g t g t g g a a a g a t g g a t a t c t a c t t a t	1260
Dd	1201	t t g a a a a a a a g c a c g a g t g g a c c t g a a g g g g t g t g g a a a g a t g g a t a t c t a c t t a t	1260
Qy	1261	a t g g a g a t c t t c t c a t g a g t t c c c a a t g c c t t c t g c g c a c g g c t c a a g c c c g	1320
Dd	1261	a t g g a g a t c t t c t c a t g a g t t c c c a a t g c c t t c t g c g c a c g g c t c a a g c c c g	1320
Qy	1321	a c g t c c t t c c a a c g c c a a c a g a t c a t a g a a c c c a a g t c g a c t t g a t c g c g a t a c a	1380
Dd	1321	a c g t c c t t c c a a c g c c c a a c a g l c a t a g a a a c c c a a g t c g a c t t g a t c g c g a t a c a	1380
Qy	1381	a t t g c a a a g t t g a g c c a g a c g c c a c g t c c g t t g a g g c g a a a t c a g a c a a g a g	1440
Dd	1381	a t t g c a a a g t t g a g c c a g a c g c c a c g t c c g t t g a g g c g a a a t c a g a c a a g a g	1440
Qy	1441	g c a t t g g t c g a t a t g a t t g c c a a g a t g a a c g a g a c a c t t g t t c c c c t t g a c g g a t c g	1500
Dd	1441	g c a t t g g t c g a t a t g a t t g c c a a g a t g a a c g a g a c a c t t g t t c c c c t t g a c g g a t c g	1500
Qy	1501	t g g t g g a c t g a g g c a a c a t c c c t g g a a g c a a c a c g t g c t t a a c c t t c a t a g g c g g g	1560
Dd	1501	t g g t g g a c t g a g g a a c a t c c c t g g a a g c a a c a c g t g c t t a a c c t t c a t a g g c g g g	1560
Qy	1561	a t g t c t c t a t g a g a g a t c t g t c a a g a a a g g t g g c a a t t g g g a t g g g t t t g a t g t g	1620
Dd	1561	a t g t c t c t a t g a g a g a t c t g t c a a g a a a g g t g g c a a t t g g g a t g g g t t t g a t g t g	1620
Qy	1621	c t t c a t g c t c c c t g c t a a	1638
Dd	1621	c t t c a t g c t c c c t g c t a a	1638

RESULT 10  
US-09-882-694A-2  
Sequence 2, Application US/09882694A  
: GENERAL INFORMATION:  
: APPLICANT: Duvick, Jon  
: APPLICANT: Maddox, Joyce  
: APPLICANT: Gilliam, Jacob  
: APPLICANT: Folkerts, Otto  
: APPLICANT: Crasta, Oswald R.  
: TITLE OF INVENTION: Compositions and Methods for Fumonisin  
: TITLE OF INVENTION: Detoxification  
: FILE REFERENCE: 5718-111  
: CURRENT APPLICATION NUMBER: US/09/882.694A  
: CURRENT FILING DATE: 2001-06-15  
: PRIOR APPLICATION NUMBER: 09/351,224

; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: flavin monooxygenase, fully spliced  
US-09-882-694A-2

Query Match 100.0%; Score 1638; DB 33; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcgccaccagcaactccagagcgattgttcctgcgcatgcgacgccaatcattg 60  
|||||  
Db 1 atgtcgccaccagcaactccagagcgattgttcctgcgcatgcgacgccaatcattg 60  
|||||  
QY 61 ggaacggcctcagcgagcatctctctgtgtacaaattgcgaaagctcagactcaagcc 120  
|||||  
Db 61 ggaacggcctcagcgagcatctctctgtgtacaaattgcgaaagctcagactcaagcc 120  
|||||  
QY 121 aaaaattcgaagagccccgattttgagcggtctggcaactggaaacgctaccctggc 180  
|||||  
Db 121 aaaaattcgaagagccccgattttgagcggtctggcaactggaaacgctaccctggc 180  
|||||  
QY 181 gctgtgtgattcggagagccctttaccactgaacttcccgagatcggaaagac 240  
|||||  
Db 181 gctgtgtgattcggagagccctttaccactgaacttcccgagatcggaaagac 240  
|||||  
QY 241 tggacctgtctgcgctatctcgaacagaaagttgtctcatatttaccactgt 300  
|||||  
Db 241 tggacctgtctgcgctatctcgaacagaaagttgtctcatatttaccactgt 300  
|||||  
QY 301 gacaagatcgggcttgagaaagactctactcgaagctgagctgattgatcgcg 360  
|||||  
Db 301 gacaagatcgggcttgagaaagactctactcgaagctgagctgattgatcgcg 360  
|||||  
QY 361 tatccagagatctgggaccttgacttcaagacgtcggtggccatttgcgacgca 420  
|||||  
Db 361 tatccagagatctgggaccttgacttcaagacgtcggtggccatttgcgacgca 420  
|||||  
QY 421 aagtatctatctcgtacggtgtgtccacaggaagcacactcccgcactcccggc 480  
|||||  
Db 421 aagtatctatctcgtacggtgtgtccacaggaagcacactcccgcactcccggc 480  
|||||  
QY 481 ctgcgcgatttcaacgggaagtgatttatttcgagtgctcgtggcagaaacttcagca 540  
|||||  
Db 481 ctgcgcgatttcaacgggaagtgatttatttcgagtgctcgtggcagaaacttcagca 540  
|||||  
QY 541 gaggccagagatcgccctatcgttcggggccacagacatccagatttcaggag 600  
|||||  
Db 541 gaggccagagatcgccctatcgttcggggccacagacatccagatttcaggag 600  
|||||  
QY 601 ttggccaaagagctacacagatcccatgtttatgcgaaggccgagctattgtctgcc 660  
|||||  
Db 601 ttggccaaagagctacacagatcccatgtttatgcgaaggccgagctattgtctgcc 660  
|||||  
QY 661 atgcggcaacgacgatgtaggaacgaacagacagccttggaggccctactaccacag 720  
|||||  
Db 661 atgcggcaacgacgatgtaggaacgaacagacagccttggaggccctactaccacag 720  
|||||  
QY 721 ctgtttgaagcagctcgaagctcgtgatttgattcccggtccagggcagctcgattggc 780  
|||||  
Db 721 ctgtttgaagcagctcgaagctcgtgatttgattcccggtccagggcagctcgattggc 780  
|||||  
QY 781 atctttgaagtcagcccccagcagcgaggcctatttcgaagagttgtggagcgtggg 840  
|||||  
Db 781 atctttgaagtcagcccccagcagcgaggcctatttcgaagagttgtggagcgtggg 840  
|||||

## RESULT 11

US-09-351-224-1  
; Sequence 1, Application US/09351224  
; GENERAL INFORMATION:  
; APPLICANT: Duwick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351,224  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1691  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera

QY 841 gcctttaattttctgttgcagtacccgagaagtcatgttgacaaaaggcccaaccga 900  
|||||  
Db 841 gcctttaattttctgttgcagtacccgagaagtcatgttgacaaaaggcccaaccga 900  
|||||  
QY 901 ctggtctatgaactctgggcccacaaagactcgtctcgtatcgtcgaatccggcacaagaga 960  
|||||  
Db 901 ctggtctatgaactctgggcccacaaagactcgtctcgtatcgtcgaatccggcacaagaga 960  
|||||  
QY 961 gatctcagctcctctggagccgactgactggttgcgtaccagcgctcccactggag 1020  
|||||  
Db 961 gatctcagctcctctggagccgactgactggttgcgtaccagcgctcccactggag 1020  
|||||  
QY 1021 agcactactacgaatactgtgacaagcgagcgctgcaaatgtgaaatctagaacaatcg 1080  
|||||  
Db 1021 agcactactacgaatactgtgacaagcgagcgctgcaaatgtgaaatctagaacaatcg 1080  
|||||  
QY 1081 cccattggctgttacaagaacaggtgtgctcttgagtgcgcaagaggggaatgc 1140  
|||||  
Db 1081 cccattggctgttacaagaacaggtgtgctcttgagtgcgcaagaggggaatgc 1140  
|||||  
QY 1141 gacacgactcgtcggcgagcggttttcgacagtttcactggctcattgacacatatgggc 1200  
|||||  
Db 1141 gacacgactcgtcggcgagcggttttcgacagtttcactggctcattgacacatatgggc 1200  
|||||  
QY 1201 ttgaaaaaacaagcacgagctgacctgaagagaggtgtgaaagatgacatctacttat 1260  
|||||  
Db 1201 ttgaaaaaacaagcacgagctgacctgaagagaggtgtgaaagatgacatctacttat 1260  
|||||  
QY 1261 atggagctctctctcatggcttccccaatgctctcttctgcgccagctcaagccccg 1320  
|||||  
Db 1261 atggagctctctctcatggcttccccaatgctctcttctgcgccagctcaagccccg 1320  
|||||  
QY 1321 accgtcttctccacggcccacgcatcatagaaccccaagtcgacttgatcgccgataca 1380  
|||||  
Db 1321 accgtcttctccacggcccacgcatcatagaaccccaagtcgacttgatcgccgataca 1380  
|||||  
QY 1381 attcgaagtgtgagcgccagcacgccctccgttgagcgacgaaatcagcacaaag 1440  
|||||  
Db 1381 attcgaagtgtgagcgccagcacgccctccgttgagcgacgaaatcagcacaaag 1440  
|||||  
QY 1441 gcatggtcgattatgattcccaagatgaacgacacactgttcccttcgacgattcg 1500  
|||||  
Db 1441 gcatggtcgattatgattcccaagatgaacgacacactgttcccttcgacgattcg 1500  
|||||  
QY 1501 tgggtgactggaggaaacatccctgggaaagcaacacgctgctttaacctcatagcgagg 1560  
|||||  
Db 1501 tgggtgactggaggaaacatccctgggaaagcaacacgctgctttaacctcatagcgagg 1560  
|||||  
QY 1561 attgctctctatgacagatctgtcaagagaggtggccaatgggattgggtttgatgtg 1620  
|||||  
Db 1561 attgctctctatgacagatctgtcaagagaggtggccaatgggattgggtttgatgtg 1620  
|||||  
QY 1621 ctctcatgctccctgctaa 1638  
|||||  
Db 1621 ctctcatgctccctgctaa 1638  
|||||

```

; FEATURE:
; OTHER INFORMATION: flavin monooxygenase with intron
US-09-351-224-1

```

Query Match 95.2%; Score 1575; DB 17; Length 1691;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 53;

[illegible]

```

RESULT 12
US 09-351-224E-1
;
; SEQUENCE INFORMATION:
; GENBANK INFORMATION:
;
; APPLICANT: Maddox, Jon
; APPLICANT: Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R
;
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; DETOXIFICATION
;
; FILE REFERENCE: 5718-111
;
; CURRENT APPLICATION NUMBER: US/09/351.224E
; CURRENT FILING DATE: 1999-07-12
;
; NUMBER OF SEQ ID NOS: 11
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 1
;
; LENGTH: 1691
;
; TYPE: DNA
;
; ORGANISM: Exophiala spinifera
;

```



Query Match 96.2%; Score 1575; DB 17; Length 1691;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 atgtcggccacacgaacatccagagcgttgcgttcgcatgagcagccatcatt 60  
DB 1 atgtcggccacacgaacatccagagcgttgcgttcgcatgagcagccatcatt 60

QY 61 gagccggccctcagcggcatctgtgtgtacaaattgcgaagctcagactaacgcc 120  
DB 61 gagccggccctcagcggcatctgtgtgtacaaattgcgaagctcagactaacgcc 120

QY 121 aaattctcggaggagcccccgtatttggcggcgtctggcactggagccgtaccctggc 180  
DB 121 aaattctcggaggagcccccgtatttggcggcgtctggcactggagccgtaccctggc 180

QY 181 gctcgtgtgtattcggagacccctctacaaactgaacttccgaagatggaagac 240  
DB 181 gctcgtgtgtattcggagacccctctacaaactgaacttccgaagatggaagac 240

QY 241 tggaaactgtgtcgcgtatctctgacacgaagagttgctgtcatatttcaacactgt 300  
DB 241 tggaaactgtgtcgcgtatctctgacacgaagagttgctgtcatatttcaacactgt 300

QY 301 gacaagatccggggtctgagaaaagacgtctacttcggagctgaggtgtgtatgcggg 360  
DB 301 gacaagatccggggtctgagaaaagacgtctacttcggagctgaggtgtgtatgcggg 360

QY 361 tatgcgaagatctgggacacttgactgtcaagacgtcggctggcgttgcgagcga 420  
DB 361 tatgcgaagatctgggacacttgactgtcaagacgtcggctggcgttgcgagcga 420

QY 421 aagatctcattctcgtactcgaaggtgtctccacaggaagacactcccgacactccggc 480  
DB 421 aagatctcattctcgtactcgaaggtgtctccacaggaagacactcccgacactccggc 480

QY 481 ctgcgcgatttcaacgggaaggtgattcattctcagtgctggcaggaagacttcgacga 540  
DB 481 ctgcgcgatttcaacgggaaggtgattcattctcagtgctggcaggaagacttcgacga 540

QY 541 gaggggccagagagtcgcgtcatcgttgcggggccacagcatccagattgttcagag 600  
DB 541 gaggggccagagagtcgcgtcatcgttgcggggccacagcatccagattgttcagag 600

QY 601 ttggccaaagaggttgaccaggtaaccattgtttatcgaggcggcgtatgtctgccc 660  
DB 601 ttggccaaagaggttgaccaggtaaccattgtttatcgaggcggcgtatgtctgccc 660

QY 661 atcgggcaacgaacgatggataggaacgaacagacgcttggaagcctactaccaccag 720  
DB 661 atcgggcaacgaacgatggataggaacgaacagacgcttggaagcctactaccaccag 720

QY 721 ctgtttgaagcagtcgaaagtctcggattggattcccggtccagcagcgtcgttggc 780  
DB 721 ctgtttgaagcagtcgaaagtctcggattggattcccggtccagcagcgtcgttggc 780

QY 781 actcttgaagtcagcccgacgcagcagcgtatcttccgaagattgtggagcgtggg 840  
DB 781 actcttgaagtcagcccgacgcagcagcgtatcttccgaagattgtggagcgtggg 840

QY 841 gctttaattttctgtctgcagtaaccgagagtcattgttgcacaaagggcccaacga 900  
DB 841 gctttaattttctgtctgcagtaaccgagagtcattgttgcacaaagggcccaacga 900

QY 901 ctggctctactgtcttgcggccaaagactcgtctcgtatcgtatccatccgcgaagaga 960  
DB 901 ctggctctactgtcttgcggccaaagactcgtctcgtatcgtatccatccgcgaagaga 960

QY 961 gatctcagtcctctcgtgagcgcctgactgttgcgtacaaagcgtcccccactgag 1020  
DB 961 gatctcagtcctctcgtgagcgcctgactgttgcgtacaaagcgtcccccactgag 1020

QY 1021 agcgactactacgaatgctggacaagccgagcgtcgaaattgtgaatctagacaacatcg 1080  
DB 1021 agcgactactacgaatgctggacaagccgagcgtcgaaattgtgaatctagacaacatcg 1080

QY 1081 cccattgtgctgttacaagaacaggtgtctcttgaatgacgacgacgaagagggaatgc 1140  
DB 1081 cccattgtgctgttacaagaacaggtgtctcttgaatgacgacgacgaagagggaatgc 1140

QY 1141 gacacgactgtctctggcagcgggtttcgaacagtttcaacttggctc----- 1184  
DB 1141 gacacgactgtctctggcagcgggtttcgaacagtttcaacttggctc----- 1184

QY 1185 -----attgacacatatgtggcttgaaaa 1207  
DB 1185 -----attgacacatatgtggcttgaaaa 1207

QY 1201 atggctccgagtcggcgttggctgaccttgaagattgacacatatgtggcttgaaaa 1260  
DB 1201 atggctccgagtcggcgttggctgaccttgaagattgacacatatgtggcttgaaaa 1260

QY 1208 acaagcacgagtggaacctgaaggaggtgtggaagatggcattacttataatggag 1267  
DB 1208 acaagcacgagtggaacctgaaggaggtgtggaagatggcattacttataatggag 1267

QY 1261 acaagcacgagtggaacctgaaggaggtgtggaagatggcattacttataatggag 1320  
DB 1261 acaagcacgagtggaacctgaaggaggtgtggaagatggcattacttataatggag 1320

QY 1268 tcttctcattgcttcccaatgccttcttcgtcgcacggtcgaagcccgaccgtcc 1327  
DB 1268 tcttctcattgcttcccaatgccttcttcgtcgcacggtcgaagcccgaccgtcc 1327

QY 1321 tcttctcattgcttcccaatgccttcttcgtcgcacggtcgaagcccgaccgtcc 1380  
DB 1321 tcttctcattgcttcccaatgccttcttcgtcgcacggtcgaagcccgaccgtcc 1380

QY 1328 ttccaacggcccaacgatcatagaaaccccaagtcgacttgatcgcgatacaaatggaa 1387  
DB 1328 ttccaacggcccaacgatcatagaaaccccaagtcgacttgatcgcgatacaaatggaa 1387

QY 1381 ttccaacggcccaacgatcatagaaaccccaagtcgacttgatcgcgatacaaatggaa 1440  
DB 1381 ttccaacggcccaacgatcatagaaaccccaagtcgacttgatcgcgatacaaatggaa 1440

QY 1388 agttgagggccgagcacccacgtccgttgaggcgacgaaatcagcacagagagcattggt 1447  
DB 1388 agttgagggccgagcacccacgtccgttgaggcgacgaaatcagcacagagagcattggt 1447

QY 1441 agttgagggccgagcacccacgtccgttgaggcgacgaaatcagcacagagagcattggt 1500  
DB 1441 agttgagggccgagcacccacgtccgttgaggcgacgaaatcagcacagagagcattggt 1500

QY 1448 gattatatttcgaagatgaacgacactcttcccttgacgagcattcgtgtgga 1507  
DB 1448 gattatatttcgaagatgaacgacactcttcccttgacgagcattcgtgtgga 1507

QY 1501 gattatatttcgaagatgaacgacactcttcccttgacgagcattcgtgtgga 1560  
DB 1501 gattatatttcgaagatgaacgacactcttcccttgacgagcattcgtgtgga 1560

QY 1508 ctggaggcaacatccctgggaagcaacacgtgttttaaccttcattaggcgggattgctc 1567  
DB 1508 ctggaggcaacatccctgggaagcaacacgtgttttaaccttcattaggcgggattgctc 1567

QY 1561 ctggaggcaacatccctgggaagcaacacgtgttttaaccttcattaggcgggattgctc 1620  
DB 1561 ctggaggcaacatccctgggaagcaacacgtgttttaaccttcattaggcgggattgctc 1620

QY 1568 tctatgacgagatctgtcgaagagaggtggcgaattgggattggattgcttccatg 1627  
DB 1568 tctatgacgagatctgtcgaagagaggtggcgaattgggattggattgcttccatg 1627

QY 1621 tctatgacgagatctgtcgaagagaggtggcgaattgggattggattgcttccatg 1680  
DB 1621 tctatgacgagatctgtcgaagagaggtggcgaattgggattggattgcttccatg 1680

QY 1628 ctccctgctaa 1638  
DB 1628 ctccctgctaa 1691

RESULT 14  
US-09-677-488-1  
; Sequence 1, Application US/09677488  
; GENERAL INFORMATION:  
; APPLICANT: Duivick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 35718/204100 (5718-111A )  
; CURRENT APPLICATION NUMBER: US/09/677,488  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1691  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: flavin monooxygenase with intron  
US-09-677-488-1

Query Match 96.2%; Score 1575; DB 26; Length 1691;

Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

```
QY 1 atgcggccaccgaacatccagagcgagatttccgtcgcatcgacgacccatcatcgtt 60
Db 1 atgcggccaccgaacatccagagcgagatttccgtcgcatcgacgacccatcatcgtt 60
QY 61 ggaagcggcctcagcggcatctctctgtgtacaaattgcgaagctcagactcaacgcc 120
Db 61 ggaagcggcctcagcggcatctctctgtgtacaaattgcgaagctcagactcaacgcc 120
QY 121 aaaaattcagagagagcccccatttggcgcgctgtgcactggaacccgtaccctggc 180
Db 121 aaaaattcagagagagcccccatttggcgcgctgtgcactggaacccgtaccctggc 180
QY 181 gctcgtgttgatcggagacgcccctctaccactgaacattcccgaaagtatgaaagac 240
Db 181 gctcgtgttgatcggagacgcccctctaccactgaacattcccgaaagtatgaaagac 240
QY 241 tggacctggcttgcgcgtatctctaccagaagagttgctgtcatatgttccaccactgt 300
Db 241 tggacctggcttgcgcgtatctctaccagaagagttgctgtcatatgttccaccactgt 300
QY 301 gacaagatccggggttgagaaaagacgtctacttcggagctgagtggttgatgcgcgg 360
Db 301 gacaagatccggggttgagaaaagacgtctacttcggagctgagtggttgatgcgcgg 360
QY 361 tatgcagagatctgggacacttgactgtcaagacgtcggtggccatgttggcagcgca 420
Db 361 tatgcagagatctgggacacttgactgtcaagacgtcggtggccatgttggcagcgca 420
QY 421 aagtatctatctcgtcactcgggtgtgtccacaggaagacacatcccgacactcccgcc 480
Db 421 aagtatctatctcgtcactcgggtgtgtccacaggaagacacatcccgacactcccgcc 480
QY 481 ctgcgcgatttcaacgggaaggtgattcttcgagtgccctggcagcagaagacttcagcga 540
Db 481 ctgcgcgatttcaacgggaaggtgattcttcgagtgccctggcagcagaagacttcagcga 540
QY 541 gaggccagagatcgccgtcatcgtgcggggccacaaagcatccagatctgttcaggag 600
Db 541 gaggccagagatcgccgtcatcgtgcggggccacaaagcatccagatctgttcaggag 600
QY 601 ttggccaaagagctgaccaggttaaccatgtttatgcgaagcccgagctattgtctgcc 660
Db 601 ttggccaaagagctgaccaggttaaccatgtttatgcgaagcccgagctattgtctgcc 660
QY 661 atgcggcaacgaacgatgataaggaacgaacagacagcgtggaagccctactaccacacg 720
Db 661 atgcggcaacgaacgatgataaggaacgaacagacagcgtggaagccctactaccacacg 720
QY 721 ctgtttgaagcagagtcgaagtcctggattggtattcccggtccagggcccggttggc 780
Db 721 ctgtttgaagcagagtcgaagtcctggattggtattcccggtccagggcccggttggc 780
QY 781 atctttgaagtcgccccagcagcgggagcgtatttcgaagagttgtggagcgtggg 840
Db 781 atctttgaagtcgccccagcagcgggagcgtatttcgaagagttgtggagcgtggg 840
QY 841 gcttttaattttctgttcggcagtagcagagagtcgtatcgtcaatcccgcaaacoga 900
Db 841 gcttttaattttctgttcggcagtagcagagagtcgtatcgtcaatcccgcaaacoga 900
QY 901 ctggtctatgactctggccaaagactcgtctcgtatcgtcaatcccgcaaacoga 960
Db 901 ctggtctatgactctggccaaagactcgtctcgtatcgtcaatcccgcaaacoga 960
QY 961 gatctcgtctctctggagcccgctactgttgggtaccgaagcgtcccccactggag 1020
Db 961 gatctcgtctctctggagcccgctactgttgggtaccgaagcgtcccccactggag 1020
QY 1021 agcactactacgaatactcgacagccgagcgtcgaaattgtgaattctagaacaatcg 1080
Db 1021 agcactactacgaatactcgacagccgagcgtcgaaattgtgaattctagaacaatcg 1080
```

```
Db 1021 agcactactacgaatactcgacagccgagcgtcgaaattgtgaattctagaacaatcg 1080
QY 1081 cccattgtgctgtttacaaagacaggtgtgctcttgaatgacgagcagcaagaggaatgc 1140
Db 1081 cccattgtgctgtttacaaagacaggtgtgctcttgaatgacgagcagcaagaggaatgc 1140
QY 1141 gacagatcgtctgtgagcaggggtttcgacaggtttcactggtcgtgagtggtctcgtc 1184
Db 1141 gacagatcgtctgtgagcaggggtttcgacaggtttcactggtcgtgagtggtctcgtc 1200
QY 1185 -----attgacacatatggcttgaaaa 1207
Db 1201 atggctcccgagtcggcagctttggctgaccttgaagagattgacacatatggcttga 1260
QY 1208 acaagcacagaatgacacctgaagaggtgtgaaagatggcatatctacttatatggag 1267
Db 1261 acaagcacggagtgacacctgaaggggtgtgaaagatggcatatctacttatatggag 1320
QY 1268 tcttctctcatggtcttcccaatgctcttctcgtcgccacggtccaagcccgacctcc 1327
Db 1321 tcttctctcatggtcttcccaatgctcttctcgtcgccacggtccaagcccgacctcc 1380
QY 1328 tttccaacggcccaacgacatcataaaccacgaagtcgactgtatcccgatacaaatgcaa 1387
Db 1381 tttccaacggcccaacgacatcataaaccacgaagtcgactgtatcccgatacaaatgcaa 1440
QY 1388 agttgagggccgagcacgcccacgtccgttgaggcgacgaaatcagcaacagagcattggt 1447
Db 1441 agttgagggccgagcacgcccacgtccgttgaggcgacgaaatcagcaacagagcattggt 1500
QY 1448 cgattatgattcccaagatgaacagcacactctgttcccttcacgagcttcggttggga 1507
Db 1501 cgattatgattcccaagatgaacagcacactctgttcccttcacgagcttcggttggga 1560
QY 1508 ctggagggcaacatccctgggaagcaacacgctgttttaaccttcataggcgagattgctc 1567
Db 1561 ctggagggcaacatccctgggaagcaacacgctgttttaaccttcataggcgagattgctc 1620
QY 1568 tctatgacagatctgtcaagagaaggtggccaattggatgggtttgatgtgcttcag 1627
Db 1621 tctatgacagatctgtcaagagaaggtggccaattggatgggtttgatgtgcttcag 1680
QY 1628 ctccctgctcaa 1638
Db 1681 ctccctgctcaa 1691
```

## RESULT 15

```
US-09-677-488a-1
; Sequence 1, Application US/09677488A
; GENERAL INFORMATION:
; APPLICANT: Duvick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; TITLE OF INVENTION: Detoxification
; FILE REFERENCE: 35718/204100
; CURRENT APPLICATION NUMBER: US/09/677,488A
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691.
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: flavin monooxygenase with intron
```





-----



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 19:25:11 ; Search time 150.7 seconds  
(without alignments)

2669,859 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638

Sequence: 1 atgtcgccaccagcaactc.....tgcttcattgctccctgctaa 1638

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	7.3	1642	US-08-232-015-2	Sequence 2, Appli
2	53.8	3.3	4403765	US-09-103-840A-2	Sequence 2, Appli
3	53.8	3.3	4411529	US-09-103-840A-1	Sequence 1, Appli
4	37.6	2.3	1365	US-09-030-995-1	Sequence 1, Appli
5	34.4	2.1	2061	US-08-472-028A-7	Sequence 7, Appli
6	34.4	2.1	2061	US-08-808-931-7	Sequence 7, Appli
7	34.4	2.1	2061	US-08-808-323-7	Sequence 7, Appli
8	34.4	2.1	2061	US-09-050-603A-7	Sequence 7, Appli
9	34.4	2.1	2061	US-09-102-420B-7	Sequence 7, Appli
10	34.4	2.1	2061	US-09-071-296-7	Sequence 7, Appli
11	34.4	2.1	2061	US-09-196-268-7	Sequence 7, Appli
12	34.4	2.1	2061	US-09-015-693-7	Sequence 7, Appli
13	34.4	2.1	2061	US-09-191-998-7	Sequence 7, Appli
14	34	2.1	2061	US-09-497-698-7	Sequence 7, Appli
15	34	2.1	2576	US-09-265-108-1	Sequence 1, Appli
16	34	2.1	2576	US-09-479-264-1	Sequence 1, Appli
17	33.8	2.1	343	US-08-473-020A-8	Sequence 8, Appli
18	33.8	2.1	2539	US-09-000-016-3	Sequence 3, Appli
19	33.8	2.1	2809	US-09-000-016-1	Sequence 1, Appli
20	33.6	2.1	779	US-08-998-416-393	Sequence 393, App
21	33.6	2.1	850	US-08-998-416-532	Sequence 532, App
22	32.8	2.0	7218	US-08-232-463-14	Sequence 14, Appli
23	32.6	2.0	28958	US-08-258-261B-6	Sequence 6, Appli
24	32.6	2.0	28958	US-08-456-837-6	Sequence 6, Appli
25	32.6	2.0	28958	US-08-457-342-6	Sequence 6, Appli
26	32.6	2.0	28958	US-08-457-646A-6	Sequence 6, Appli
27	32.6	2.0	28958	US-08-458-076A-6	Sequence 6, Appli

28	32.6	2.0	28958	1	US-08-764-233A-4	Sequence 4, Appli
29	32.6	2.0	28958	1	US-08-457-335A-6	Sequence 6, Appli
30	32.6	2.0	28958	1	US-08-729-214-6	Sequence 6, Appli
31	32.6	2.0	28958	3	US-09-028-934-6	Sequence 6, Appli
32	32.6	2.0	49377	1	US-08-764-233A-1	Sequence 1, Appli
33	32.2	2.0	288	2	US-08-716-942-16	Sequence 16, Appli
34	32.2	2.0	30001	1	US-08-125-468-1	Sequence 1, Appli
35	32.2	2.0	30001	2	US-08-474-933-1	Sequence 1, Appli
36	32	2.0	23673	4	US-09-773-816-1	Sequence 3, Appli
37	31.8	1.9	1498	4	US-09-461-474-3	Sequence 36, Appli
38	31.8	1.9	4041	4	US-09-105-537-36	Sequence 1, Appli
39	31.8	1.9	36778	4	US-09-105-537-5	Sequence 3, Appli
40	31.8	1.9	38506	3	US-09-320-878-19	Sequence 5, Appli
41	31.6	1.9	1164	2	US-08-997-080-193	Sequence 193, App
42	31.6	1.9	1164	2	US-08-997-362-193	Sequence 193, App
43	31.6	1.9	1164	4	US-09-095-855-193	Sequence 193, App
44	31.6	1.9	1164	4	US-09-324-542-193	Sequence 193, App
45	31.2	1.9	249	1	US-08-482-385A-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-08-232-015-2  
Sequence 2, Application US/08232015  
Patent No. 5650272  
GENERAL INFORMATION:  
APPLICANT: GUESDON, Jean-Luc  
APPLICANT: THIERRY, Veronique  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH HYBRIDIZE SPECIFICALLY WITH BACTERIAL STRAINS OF THE MYCOBACTERIUM AVIUM-INTRACELLULAR COMPLEX  
TITLE OF INVENTION: AVIUM-INTRACELLULAR COMPLEX  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/232,015  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 93/09251  
FILING DATE: 13-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 91/13504  
FILING DATE: 31-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Craze-Feury, Sharon E  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 004900-127  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1642 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO





```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,028A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1748/CIP
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 64..1698
OTHER INFORMATION: /note= "Maize protox-2 cDNA;
US-08-472-028A-7

Query Match 2.1%; Score 34.4; DB 1; Length 2061;
Best Local Similarity 59.0%; Pred. No. 1.9;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 39 cgcagtcgacccatcatcttgagcgcgcctcagcgcctcctctgtgtacaaatt 98
Db 195 CGCCAGATCGTCGCGCTCGCGCGCGGCTCAGCGGCTCGCGCGGCTACAGGCT 254

Qy 99 gcgaagctcagactcaacgccaaatcttcgaggagcc 138
Db 255 CAGACAGAGCGCGGTGAACGTAAACGGTGTTCGAAGCGGCC 294

RESULT 6
US-08-808-931-7
; Sequence 7, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Heifetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA

```

```

ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
OTHER INFORMATION: /note= "Maize protox-2"
US-08-808-931-7

Query Match 2.1%; Score 34.4; DB 2; Length 2061;
Best Local Similarity 59.0%; Pred. No. 1.9;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 39 cgcagtcgacccatcatcttgagcgcgcctcagcgcctcctctgtgtacaaatt 98
Db 195 CGCCAGATCGTCGCGCTCGCGCGCGGCTCAGCGGCTCGCGCGGCTACAGGCT 254

Qy 99 gcgaagctcagactcaacgccaaatcttcgaggagcc 138
Db 255 CAGACAGAGCGCGGTGAACGTAAACGGTGTTCGAAGCGGCC 294

RESULT 7
US-08-808-323-7
; Sequence 7, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; APPLICANT: Heifetz, Peter
; TITLE OF INVENTION: Promoters from Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018105artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005

```









TITLE OF INVENTION: Oxidase Enzyme Activity In Eukaryotic Organisms  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/09/191,998  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/261,198  
FILING DATE: 16-JUN-94  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1748/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 64..1698  
OTHER INFORMATION: /note= "Maize protox-2 cDNA;  
OTHER INFORMATION: sequence from pWDC-3"  
US-09-191-998-7  
Query Match 2.1%; Score 34.4; DB 4; Length 2061;  
Best Local Similarity 59.0%; Pred. No. 1.9;  
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
Qy 39 cgcagcgagccatcattgttgagccgctcagcgccatctctgtgtgtacaaatt 98  
Db 195 CGCAGATCGTCCGCTGCTGCGCGCGGCTCAGCGGCTCGCGGGCGGTACAGGCT 254  
Qy 99 gcgaagctcagactcaagcgcacaaattcttcgagggagcc 138  
Db 255 CAGACAGAGCGCGTGAACGTACGCGTGTTCGAAGCGCC 294  
RESULT 14  
US-09-497-698-7  
Sequence 7, Application US/09497698  
Patent No. 6308458  
GENERAL INFORMATION:  
APPLICANT: Volrath, Sandra  
Johnson, Marie  
Ward, Eric  
Helretz, Peter  
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN  
OXIDASE ("PROTOX")  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6308458artis Corporation  
STREET: 3054 Cornwallis Road

CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/09/497,698  
FILING DATE: 03-Feb-2000  
CLASSIFICATION: <Unknown>  
30-MAR-1998  
11-MAR-1998  
28-FEB-1997  
28-FEB-1996  
21-JUN-1996  
06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/102,420  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 09/050,603  
FILING DATE: 30-MAR-1998  
APPLICATION NUMBER: US 60/126,430  
FILING DATE: 11-MAR-1998  
APPLICATION NUMBER: US 08/808,931  
FILING DATE: 28-FEB-1997  
APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
APPLICATION NUMBER: US 08/472,028  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Zea mays (maize)  
ORGANISM: Zea mays (maize)  
IMMEDIATE SOURCE:  
CLONE: pWDC-3 (NRRL B-21259)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 64..1698  
OTHER INFORMATION: /product= "Maize protox-2"  
SEQUENCE DESCRIPTION: SEQ ID NO: 7;  
US-09-497-698-7  
Query Match 2.1%; Score 34.4; DB 4; Length 2061;  
Best Local Similarity 59.0%; Pred. No. 1.9;  
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
Qy 39 cgcagcgagccatcattgttgagccgctcagcgccatctctgtgtgtacaaatt 98  
Db 195 CGCAGATCGTCCGCTGCTGCGCGCGGCTCAGCGGCTCGCGGGCGGTACAGGCT 254





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: April 26, 2002, 19:26:06 ; Search time 706.01 Seconds  
(without alignments)  
3983.380 Million cell updates/sec

Title: US-09-882-694A-2  
Perfect score: 1638  
Sequence: 1 atgcgcgcaccagcaactc.....tgcttcattgcctcctgctaa 1638

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1638	100.0	1638	22	AAD02691
2	1575	96.2	1691	22	AAD02690
3	153.2	9.7	1695	21	AAF12444
4	119	7.3	1642	14	AAQ41978
5	80.8	4.9	702	21	AAF12000
c 6	53.8	3.3	4403765	22	AAI99683
c 7	53.8	3.3	4411529	22	AAI99682
8	46.6	2.8	1812	22	AAH22871
9	46.6	2.8	10480	22	AAH22877

c 10	41.8	2.6	534720	19	AAV30458
c 11	41.8	2.6	536165	19	AAV30459
12	41.4	2.5	10524	23	AAS59593
13	40.2	2.5	1677	21	AAS91256
14	40.2	2.5	24494	21	AAS91253
15	37.6	2.3	747	22	AAH76382
16	37.6	2.3	1365	20	AAH21172
17	37.6	2.3	1812	22	AAH22890
18	36.8	2.2	362	21	AAH15987
19	36.6	2.2	733	21	AAH62333
20	35.6	2.2	1812	22	AAH22888
21	35.4	2.2	4590	22	AAH24065
22	35.2	2.1	536	21	AAH291289
c 23	35	2.1	2132	23	AAS68025
c 24	34.8	2.1	1769	23	ABL12489
c 25	34.8	2.1	3769	23	ABL12488
c 26	34.8	2.1	6077	23	ABL12266
27	34.4	2.1	265	21	AAH10532
28	34.4	2.1	356	21	AAH31895
29	34.4	2.1	2051	22	AAH15430
30	34.4	2.1	2061	17	AAH11676
31	34.4	2.1	2061	18	AAV04307
32	34.4	2.1	2061	18	AAT86131
33	34.4	2.1	2061	19	AAV07254
34	34.4	2.1	2061	22	AAD21860
35	34.4	2.1	2061	22	AAF76574
36	34.4	2.1	2061	22	AAD17334
37	34	2.1	435	21	AAF08106
38	34	2.1	2576	21	AAH295701
39	34	2.1	2576	22	AAH47743
c 40	34	2.1	4216	8	AAH70559
41	33.8	2.1	343	11	AAH06511
c 42	33.8	2.1	936	22	AAH66000
43	33.8	2.1	1647	22	AAH66001
c 44	33.8	2.1	2539	18	AAH61454
45	33.8	2.1	2809	18	AAH61455

## ALIGNMENTS

RESULT 1

AAD02691  
ID AAD02691 standard; DNA; 1638 BP.

XX AAD02691;

DT 02-MAY-2001 (first entry)

XX Exophiala spinifera flavin monooxygenase coding sequence.

XX Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
XX aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;  
XX detoxification; mycotoxin; animal feed; human feed; silage;  
XX transgenic plant; transgenic animal; microbial spray; ds.

XX Exophiala spinifera.

XX Key Location/Qualifiers

FT CDS 1..1638

FT /tag= a

FT /product= "E. spinifera flavin monooxygenase"

XX MO200105980-A1.

XX 25-JAN-2001.

XX 14-JUL-1999; 99WO-US15824.

XX 14-JUL-1999; 99WO-US15824.

XX (PION-) PIONEER HI-BRED INT INC.  
XX (CURA-) CURAGEN CORP.

Rhizobium species  
Rhizobium species  
Propionibacterium  
Flavin monooxygenase  
Bacterium 2412.1 f  
Nucleotide sequenc  
N. asteroides stra  
R. ruber cyclodode  
Plant microsatelli  
Human Mad3 gene se  
R. ruber cyclodode  
Yeast AAD9604-asso  
Fumonisin-cataboli  
DNA encoding novel  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Human secreted pro  
Plant microsatelli  
Maize protoporphyr  
DNA encoding a pro  
Maize prototox-2 cDN  
Maize protoporphyr  
Maize cDNA encodin  
Maize protoporphyr  
Maize protoporphyr  
Maize protoporphyr  
Fusarium venenatum  
Humicola insolens  
H. insolens DSM 18  
Sequence of nitrog  
Encodes protein XX  
C glutamicum codin  
C glutamicum codin  
Streptomyces virid  
Dhpa-mel chimeric



Db 1561 attctctcatgagcagatcctcgaagagaggtggccaaattggagtggttgatgtg 1620  
|||||  
QY 1621 attcatgctccctgctaa 1638  
|||||  
Db 1621 attcatgctccctgctaa 1638  
|||||  
RESULT 2  
ID AAD02690  
AC AAD02690 standard; DNA; 1691 BP.  
XX AAD02690;  
XX  
XX  
DT 02-MAY-2001 (first entry)  
DE Exophiala spinifera flavin monooxygenase genomic DNA.  
XX  
XX Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
KW aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;  
KW detoxification; mycotoxin; animal feed; human feed; silage;  
KW transgenic plant; transgenic animal; microbial spray; ds.  
XX  
XX Exophiala spinifera.  
OS  
XX  
XX WO200105980-A1.  
PN  
XX  
XX 25-JAN-2001.  
PD  
XX  
XX 14-JUL-1999; 99WO-US15824.  
XX  
XX 14-JUL-1999; 99WO-US15824.  
XX  
XX {PUN-} PIONEER HI-BRED INT INC.  
PA {CURA-} CURAGEN CORP.  
PI  
PI Duwick JP, Maddox J, Gilliam J, Folkerts O, Crasta OR;  
DR  
DR WPI; 2001-147345/15.  
DR P-PSDB; AAY72634.  
XX  
XX Novel polynucleotides encoding Exophiala degradative or transport  
PT enzyme which is useful for detoxifying fumonisin or structurally  
PT related mycotoxin during processing of grain for human or animal food  
PT consumption  
XX  
XX Claim 1b; Page 63-64; 90pp; English.  
XX  
XX The patent discloses novel polynucleotides encoding Exophiala spinifera  
CC fumonisin degradative or transport enzymes such as flavin monooxygenase,  
CC aldehyde dehydrogenase, permease and p-glycoprotein. A fumonisin esterase  
CC enzyme with at least one fumonisin degradative enzyme is useful for  
CC detoxifying fumonisin or a structurally related mycotoxin present in  
CC harvested grain, on application to a plant or to harvested grain  
CC during processing, or to processed grain that is to be used as animal  
CC or human feed, or as a silage. Nucleotide fragments of the present  
CC invention are useful as probes and primers. They can be introduced  
CC into microorganisms that multiply on plants to deliver enzymes to  
CC potential target crops. The genes encoding the degrading enzymes are  
CC introduced via a vector into a microbial host and the transformed host  
CC is supplied to the environment, plants or animals for reducing the  
CC pathogenicity of a fungus producing fumonisin. The genes of the  
CC invention are fermented in a bacterial host and the resulting bacteria  
CC is processed and used as a microbial spray. The nucleotide sequences  
CC can be used alone or in combination to engineer microbes or other  
CC organisms to metabolise fumonisin and resist its toxic effects.  
CC The present sequence is a genomic DNA encoding Exophiala spinifera  
CC flavin monooxygenase, a fumonisin degradative enzyme.  
XX  
XX Sequence 1691 BP; 397 A; 444 C; 480 G; 370 T; 0 other;

Query Match 96.2%; Score 1575; DB 22; Length 1691;

Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 53; Gaps 1;  
QY 1 atgtcgccaccagcaactccagagcgattgttcgctcgatcgacgccaatcattgtt 60  
|||||  
Db 1 atgtcgccaccagcaactccagagcgattgttcgctcgatcgacgccaatcattgtt 60  
|||||  
QY 61 ggagcgccctcagcgccatctctgctgtgtacaaattcgaaagctcagactcaagcc 120  
|||||  
Db 61 ggagcgccctcagcgccatctctgctgtgtacaaattcgaaagctcagactcaagcc 120  
|||||  
QY 121 aaaatctcgaggagcccccgattttggcgctgtgctgactggaacccctaccctggc 180  
|||||  
Db 121 aaaatctcgaggagcccccgattttggcgctgtgctgactggaacccctaccctggc 180  
|||||  
QY 181 gctcggttgattcgagagccctctaccgaactgaacattcccggaagtatggaagac 240  
|||||  
Db 181 gctcggttgattcgagagccctctaccgaactgaacattcccggaagtatggaagac 240  
|||||  
QY 241 tggacctgtcttcgcgtatccctgacacagaagattgtctgcatatgttcacactgt 300  
|||||  
Db 241 tggacctgtcttcgcgtatccctgacacagaagattgtctgcatatgttcacactgt 300  
|||||  
QY 301 gacaagatccggggcttgagaaagactctacttcgagactgagtggttgatgcgcgc 360  
|||||  
Db 301 gacaagatccggggcttgagaaagactctacttcgagactgagtggttgatgcgcgc 360  
|||||  
QY 361 tatgccagagatctgggcaactggactgtcaagactcggtggccattgttgacagacca 420  
|||||  
Db 361 tatgccagagatctgggcaactggactgtcaagactcggtggccattgttgacagacca 420  
|||||  
QY 421 aagtattctattctcgctacgctgggtgtctccacaggaagcacactcccgactcccgcc 480  
|||||  
Db 421 aagtattctattctcgctacgctgggtgtctccacaggaagcacactcccgactcccgcc 480  
|||||  
QY 481 ctgcgcgatttcaacgggaaggtgattcattcagtcgctggcgacagagacttcgacgca 540  
|||||  
Db 481 ctgcgcgatttcaacgggaaggtgattcattcagtcgctggcgacagagacttcgacgca 540  
|||||  
QY 541 gagggccagagatcgccgctcatcgtgctggggccacacagcatccagattgttcaggag 600  
|||||  
Db 541 gagggccagagatcgccgctcatcgtgctggggccacacagcatccagattgttcaggag 600  
|||||  
QY 601 ttggccaagaagctgacacaggttaaccattgttatcggaaggccgactattgtctccc 660  
|||||  
Db 601 ttggccaagaagctgacacaggttaaccattgttatcggaaggccgactattgtctccc 660  
|||||  
QY 661 atcgcccaacgaacatgatatgatacgaacacacagcagcctggaagcctactaccacac 720  
|||||  
Db 661 atcgcccaacgaacatgatatgatacgaacacacagcagcctggaagcctactaccacac 720  
|||||  
QY 721 ctgtttgaagcagatcgaaagctcggattcgattcccggtccagggaccctcggttgcc 780  
|||||  
Db 721 ctgtttgaagcagatcgaaagctcggattcgattcccggtccagggaccctcggttgcc 780  
|||||  
QY 781 atctttgaagtcagcccccgacgagcgaggagcctatttcgaagagttgtgggagcgtggg 840  
|||||  
Db 781 atctttgaagtcagcccccgacgagcgaggagcctatttcgaagagttgtgggagcgtggg 840  
|||||  
QY 841 gccttaattctctgcttgcagttaccgagagtcattggttgacaaaagcccaaccca 900  
|||||  
Db 841 gccttaattctctgcttgcagttaccgagagtcattggttgacaaaagcccaaccca 900  
|||||  
QY 901 ctggtctatgacttctggggccaaaagactcgtatcctgctatcgtcaatccggcaagaga 960  
|||||  
Db 901 ctggtctatgacttctggggccaaaagactcgtatcctgctatcgtcaatccggcaagaga 960  
|||||  
QY 961 gatctcatggtcctctggagcccgctactggttcggtacccaagcgtcccccactggag 1020  
|||||  
Db 961 gatctcatggtcctctggagcccgctactggttcggtacccaagcgtcccccactggag 1020  
|||||  
QY 1021 agggactactaagaatactgacacacgagcgctgaaattgtgaatctgaagaacatcg 1080  
|||||









DR WPI: 2001-647261/74.  
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises  
PT determining the nucleotide sequence of the strain at positions in the  
PT genome corresponding to positions where M. tuberculosis strains CDC  
PT 1551 and H37Rv differ  
XX  
XX Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.  
PS  
XX The invention relates to evaluating strain variation within and between  
CC different populations of the tuberculosis bacterial pathogen,  
CC Mycobacterium tuberculosis or related Mycobacterium by determining the  
CC nucleotide sequence of the first strain at positions in the complete  
CC sequence of the genome that correspond to positions that differ in the  
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and  
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of  
CC M. tuberculosis and has valuable application in the fields of  
CC tuberculosis genetics, epidemiology, patient treatment and epidemic  
CC monitoring.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.  
XX  
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

Query Match 3.3%; Score 53.8; DB 22; Length 4403765;  
Best Local Similarity 49.0%; Pred. No. 0.00049;  
Matches 307; Conservative 0; Mismatches 307; Indels 12; Gaps 2;  
QY 39 cgcacgacgccatctctgttgagcgccctcagggccatctctgtgtgacaaatt 98  
Db 1569416 CGACTACACGCACTACGTCGCGCGGGTTCTCGGAGTCGGCGCAATCAGCT 1569357  
QY 99 ggaagctcagactcagcccaaatcttcgagggagcccccatttggcggggt--- 155  
Db 1569356 CGACCGGGCGGGATTTCAGCGACTACCTGCTGTCGAAGCGCGGCTGTGGCGGGAC 1569297  
QY 156 ctggcactggaaacgcctacccctggcctctgttgatcgagagccctcttaccact 215  
Db 1569296 CTGGCACTGAACACACTATCCCGGTATCGCGGTGATATTCCTCTCTATCAGTT 1569237  
QY 216 gaacattccggaagtatgaaagactggactctgtctgtcgcctctctcaccagaga 275  
Db 1569236 -----CTCTTCGACGACGCGGCACTGTGCGGCACTATGACCGGACGCA 1569186  
QY 276 gttgctgtcatatgttcaacactgtgaaagatccgggcttgaaagagactctactt 335  
Db 1569185 GCTGAAGGCTTACCGCAACACTGCGTGCAGAAATACGGATCCGGTCCGGATCCGGTT 1569126  
QY 336 cggagctgagtggtgatgcgcggtatgcaagagatctgggaacctggactgtcaagac 395  
Db 1569125 GAACACCAAGGTGCTTGGCGCGGAGTTCGATGACGACAGCCTGTGGCGGGTCAAC 1569066  
QY 396 gtcggctgacctatgttcgacggcaagatctctcattctcgtacggggtgtccacag 455  
Db 1569065 GCACCGCGGGCGGAAATACACAGCAGGTTCTCTGATCAGCGCATTCGACGGT 1569006  
QY 456 gaagcacactcccgactcccgctcgcggtatctcaacggaggaaggtattctcag 515  
Db 1569005 GCCCAAACTACCGACATCGACGGGTGACTCTGTTGAAGGGGTCCACATGCACACCG 1568946  
QY 516 tgcctgcacgaagacttcgacgagagagagagagactcgcgtcagcgtggtcggggc 575  
Db 1568945 ACGCTGGGATACACACCGACCTGACCGGACGCGGTGCGGATCTTCGCGCCGGGCG 1568886  
QY 576 caacagactccagattgttcagagcttgcccaagaagctgacacgaactatgttat 635  
Db 1568886 CTCGGCGGTGCAAGTCAATTCGGAGATGGCGCAATTTGTCACCTCAGCGTTTTC 1568826  
QY 636 gcgaaggccgagctattgtctgccc 661  
Db 1568825 GCGACACCCCTATTGTGTGTTCCTCCCA 1568800

RESULT 7  
AAI99682/C  
ID AAI99682 standard; DNA; 4411529 BP.  
XX  
AC AAI99682;  
XX  
XX 15-JAN-2002 (first entry)  
XX  
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.  
XX  
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;  
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
XX US6294328-B1.  
XX  
XX 25-SEP-2001.  
XX  
PF 24-JUN-1998; 98US-0103840.  
XX  
PR 24-JUN-1998; 98US-0103840.  
XX  
XX (GENO-) INST GENOMIC RES.  
XX  
PI Fleischmann RD, White OR, Fraser CM, Venter JC;  
XX  
XX WPI: 2001-647261/74.  
XX  
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises  
PT determining the nucleotide sequence of the strain at positions in the  
PT genome corresponding to positions where M. tuberculosis strains CDC  
PT 1551 and H37Rv differ  
XX  
XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.  
XX  
XX The invention relates to evaluating strain variation within and between  
CC different populations of the tuberculosis bacterial pathogen,  
CC Mycobacterium tuberculosis or related Mycobacterium by determining the  
CC nucleotide sequence of the first strain at positions in the complete  
CC sequence of the genome that correspond to positions that differ in the  
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and  
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of  
CC M. tuberculosis and has valuable application in the fields of  
CC tuberculosis genetics, epidemiology, patient treatment and epidemic  
CC monitoring.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.  
XX  
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Query Match 3.3%; Score 53.8; DB 22; Length 4411529;  
Best Local Similarity 49.0%; Pred. No. 0.00049;  
Matches 307; Conservative 0; Mismatches 307; Indels 12; Gaps 2;  
QY 39 cgcacgacgccatctctgttgagcgccctcagggccatctctgtgtgacaaatt 98  
Db 1569377 CGACTACACGCACTACGTCGCGCGGGTTCTCGGAGTCGGCGCAATCAGCT 1569318  
QY 99 gcaaaactcagactcagcccaaatcttcgagggagcccccatttggcggggt--- 155  
Db 1569317 CGACCGGGCGGGATTTCAGCGACTACCTGCTGTCGAAGCGCGGAGTCGCGGGAC 1569158  
QY 156 ctggcactggaaacgcctacccctggcctctgttgatcgagagccctcttaccact 215  
Db 1569157 CTGGCACTGAACACACTATCCCGGTATCGCGGTGATATTCCTCTCTATCAGTT 1569398  
QY 216 gaacattcccggaagtatgaaagactggactgctcttgcgcgtctcctcagcagaaga 275

Db 1569397 -----CFTCTTGACGAGCGCGCACTGTGTCGGCACCCTATGCACCCGGACAGCA 1569347  
 Qy 276 gttgctgtcatatgttccacactgtgacaagatccggggcttgagaaagacgtctactt 335  
 Db 1569346 GCTGAAGCCCTACCGCAACACTGCTGCACAAATACGGCATCCGGTTCGGGATCCGGTT 1569287  
 Qy 336 cggagctaggtggttgaatcgcggtatccagagatctggcaccctggactgcaagac 395  
 Db 1569286 GAACACCAAGTCTTCCGCCGAGTTCGATGACGACACAGCCTGTGGCGGTCGAAAC 1569227  
 Qy 396 gtccgtgagccatctgcgaagcaagatattctcattctcgtcaggggtgtgtccacag 455  
 Db 1569226 GAACCCGGCGCGGAAATACACACAGGTTCTGATACGCCATCGCGCATCTGACGGT 1569167  
 Qy 456 gaagacacctccgcacactcccccgcgtccgatttccaaacgggaaggtgatttcattcgag 515  
 Db 1569166 GCCCAACTACCCGACATCGACGGGTGGACTCGTTTGAAGGGGTWCACCATGCACACCGC 1569107  
 Qy 516 tgctggcaggaacttcgacgagagggccagagagtcgcgcgtcatcgtggtccggggc 575  
 Db 1569106 AGCTGGGATCACACCCAGGACCTGACCGGCAAGCGGTCCGGATCATCGGACCGGGGC 1569047  
 Qy 576 cacaagcatccagattgttcaggagttggccaagaaggctgaccaggttaaccattgttat 635  
 Db 1569046 CTCGCGGTGAGGTATTCGCGAGATGGCGCAATTGTCTCCACCTCACCGTTTTC 1568987  
 Qy 636 gcaagggccgagctattgtctgccca 661  
 Db 1568986 GCGCACCCCTATTGTGTGTTCCCA 1568961

## RESULT 8

AAH22871  
 ID AAH22871 standard; DNA; 1812 BP.

AC AAH22871;

DT 17-SEP-2001 (first entry)

DE R. ruber cdda gene ORF encoding cyclododecanone monooxygenase.

XX Cyclododecanone; degradation: lauryl lactone esterase; cdda; cddx; cddx;  
 KW dodecanoic diacid; cyclododecanone monooxygenase; cyclic ketone; cddy;  
 KW 12-hydroxylauric acid dehydrogenase; 12-oxo lauric acid dehydrogenase;  
 KW macrolactone; microbiological; cddc; cddb; ds.

XX Rhodococcus ruber.

XX Key Location/Qualifiers

PH CDS 1..1812

FT /\*tag=

FT /gene= "cdda"

FT /product= "cyclododecanone monooxygenase"

XX WQ200142436-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000MO-US33426.

XX PR 10-DEC-1999; 99US-0170214.

XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Chen MW, Cheng Q, Gibson KJ, Kostichka KN, Thomas SM;

XX WPI; 2001-451630/48.

XX P-PSDB; AAB85234.

XX Isolated nucleic acids encoding dodecanoic diacid synthesizing enzyme,  
 PT cyclododecanone monooxygenase for bioproduction of dodecanoic diacid  
 PT from cyclododecanone

PS Claim 8; Page 51; 78pp; English.

XX The invention relates to genes from Rhodococcus ruber involved in  
 CC cyclododecanone degradation pathway. A 10 kb cluster from R. ruber was  
 CC isolated that comprises genes (designated cdda, cddb, cddx, cddy, cddc  
 CC and cddp) encoding enzymes such as dodecanoic diacid synthesizing enzyme,  
 CC cyclododecanone monooxygenase, lauryl lactone esterase, 12-hydroxylauric  
 CC acid dehydrogenase, 12-oxo lauric acid dehydrogenase. Host cells  
 CC comprising the nucleic acids are used for the production of dodecanedioic  
 CC acid when contacted with cyclododecanone, lauryl lactone when contacted  
 CC with cyclododecanone, 12-hydroxy lauric acid when contacted with lauryl  
 CC lactone, 12-oxo lauric acid when contacted with 12-hydroxy lauric acid  
 CC and dodecanedioic acid when contacted with 12-oxo lauric acid. They  
 CC are also useful for the production of hydroxy acids when contacted with  
 CC cyclic ketones with 6 to eight carbon atoms like C6, C10, C11, C12, C13  
 CC and C15 cyclic ketones, cyclohexanone, cyclodecanone, cycloundecanone,  
 CC cyclo dodecanone, cyclotridecanone and cyclotetradecanone. The host cells  
 CC are also used for the production of macrolactones with at least 10 carbon  
 CC atoms when contacted with a cyclic ketone. The nucleic acids are useful  
 CC for the bioproduction of dodecanoic diacid from cyclododecanone by  
 CC microbiological means. The present sequence represents the nucleotide  
 CC sequence of ORF1 (cdda gene) encoding a cyclododecanone monooxygenase  
 CC enzyme isolated from a 10kb nucleic acid fragment from R. ruber SCl.  
 XX  
 SQ Sequence 1812 BP; 368 A; 598 C; 571 G; 275 T; 0 other;

Query Match 2.8%; Score 46.6; DB 22; Length 1812;

Best Local Similarity 45.7%; Pred. No. 0.0012;

Matches 206; Conservative 0; Mismatches 239; Indels 6; Gaps 1;

QY 115 aacgccaaaatcttcagggagagcccgatcttggcggcgtctggaactggaaccgctac 174

DB 244 agcgtccgcatcatcgacagggccggcacttcggcgtcgtgctgactggaacagatac 303

QY 175 cctggcgtcgtgtgtgattcggagacgcccctctctaccgaacttcccgaagtatgg 234

DB 304 ccggcgcgatgtgtgacacgcgagccatggtgtacatgcgcgctgctcgaggaaaccgac 363

QY 235 aaagactggaactggtcttgcgcctatcctgaccagaaagagttgctgtcatatgttccac 294

DB 364 tacatgccaccagaga-----agtagcgcgacgcccggagagactcctcgacactgccaa 417

QY 295 cactgtgacaagatccggggttgagaaagacgtctacttcgagagctgaggtggttgat 354

DB 418 cgaatcgggcaaacactacgactgtgacgacgacgctgttccacacgcgaagtcacgcac 477

QY 355 gcgcggtatgccagagatctgggcacactgtcaagactcgctggccatgttgcg 414

DB 478 ctggtctggcaggagacagatcagcgtggtggtatctcgacgaaccgaggtgaccacttc 537

QY 415 acggcaaaagtatctcattctcgtacaggggttgcacacaggaagacacactccgcaetc 474

DB 538 acggcccaattcgtgggtatgggacacggccctctgacgtggcgacgtcgccggcatc 597

QY 475 cccggcctcgccgatttccaaagggaagtgattcattctcgagctgctggcagcgaagacttc 534

DB 598 cccgggagtcgaatcttcgcggcaagtcgttccacaccagtcgattgggactacgactac 657

QY 535 gacgcagagggccagagatcgccgtcatcg 565

DB 658 accggcgcgacgctggcgccgcatgg 688

## RESULT 9

AAH22877

ID AAH22877 standard; DNA; 10480 BP.

XX AAH22877;

XX 17-SEP-2001 (first entry)

XX R. ruber 10kb gene cluster nucleotide sequence.



```
FT FT /product= "glutamate dehydrogenase-like protein"
FT FT /note= "homologous to the GLUD1 gene"
FT FT /tag= i
FT CDS complement (430538..431284)
FT FT
FT FT /standard_name= "ORF K9"
FT FT /product= "transposase homologue"
FT FT /tag= j
FT CDS complement (431296..432840)
FT FT
FT FT /standard_name= "ORF K10"
FT FT /product= "transposase homologue"
FT FT /note= "homologous to the Tnp gene"
FT FT /tag= k
FT CDS complement (433880..434110)
FT FT
FT FT /standard_name= "ORF K11"
FT FT /product= "protein of unknown function"
FT FT /note= "homologous to the FixU gene"
FT CDS complement (434107..434433)
FT FT
FT FT /standard_name= "ORF K12"
FT FT /product= "protein of unknown function"
FT FT /tag= m
FT CDS complement (434517..434711)
FT FT
FT FT /standard_name= "ORF K13"
FT FT /product= "ferrodoxin/ferrodoxin-like protein"
FT FT /note= "homologous to the FdxN gene"
FT CDS complement (434753..436234)
FT FT
FT FT /tag= n
FT FT /standard_name= "ORF K14"
FT FT /gene= "nifB"
FT FT /product= "protein involved in FeMo co-factor
FT FT biosynthesis"
FT CDS complement (436460..438130)
FT FT
FT FT /tag= o
FT FT /standard_name= "ORF K15"
FT FT /gene= "nifA"
FT FT /product= "positive regulator of nif, fix and other
FT FT genes"
FT CDS complement (438297..438590)
FT FT
FT FT /tag= p
FT FT /standard_name= "ORF K16"
FT FT /gene= "fixX"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (438605..439912)
FT FT
FT FT /tag= q
FT FT /standard_name= "ORF K17"
FT FT /gene= "fixC"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (439923..441032)
FT FT
FT FT /tag= r
FT FT /standard_name= "ORF K18"
FT FT /gene= "fixB"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (441042..441899)
FT FT
FT FT /tag= s
FT FT /standard_name= "ORF K19"
FT FT /gene= "fixA"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (442316..442636)
FT FT
FT FT /tag= t
FT FT /standard_name= "ORF K20"
FT FT /product= "protein of unknown function"
FT CDS complement (443313..443879)
FT FT
FT FT /tag= u
FT FT /standard_name= "ORF K21"
FT FT /product= "protein of unknown function"
FT CDS 444337..445029
FT FT
FT FT /tag= v
FT FT /standard_name= "ORF K22"
FT FT /product= "ferrodoxin-like protein"
FT FT /note= "homologous to the NifQ gene"
FT CDS 445088..446602
FT FT
FT FT /tag= w
FT FT /standard_name= "ORF K23"
```

```
FT FT /gene= "dctA"
FT FT /product= "C4-dicarboxylate transport protein"
FT FT /note= "homologous to the DctAI gene"
FT CDS 446599..447843
FT FT
FT FT /tag= x
FT FT /standard_name= "ORF L1"
FT FT /product= "cytochrome P450-like protein"
FT FT /note= "homologous to the CamC gene"
FT CDS 447844..448500
FT FT
FT FT /tag= y
FT FT /standard_name= "ORF L2"
FT FT /product= "gamma-hexachlorocyclohexane-dechlorinase-like
FT FT protein"
FT CDS 448497..450203
FT FT
FT FT /tag= z
FT FT /standard_name= "ORF L3"
FT FT /product= "putative protein with degradative function"
FT CDS 450341..451396
FT FT
FT FT /tag= aa
FT FT /standard_name= "ORF L4"
FT FT /product= "luciferase alpha-subunit-like protein"
FT FT /note= "homologous to the luxA gene"
FT CDS 452980..454494
FT FT
FT FT /tag= ab
FT FT /standard_name= "ORF L6"
FT FT /gene= "nifD"
FT FT /product= "alpha-subunit of FeMo protein of nitrogenase"
FT CDS 454590..456131
FT FT
FT FT /tag= ac
FT FT /standard_name= "ORF L7"
FT FT /gene= "nifK"
FT FT /product= "beta-subunit of FeMo protein of nitrogenase"
FT CDS 456187..457677
FT FT
FT FT /tag= ad
FT FT /standard_name= "ORF L8"
FT FT /product= "protein involved in FeMo co-factor
FT FT biosynthesis"
FT CDS 457687..459096
FT FT
FT FT /tag= ae
FT FT /standard_name= "ORF L9"
FT FT /product= "protein involved in FeMo co-factor
FT FT biosynthesis"
FT CDS 459093..459575
FT FT
FT FT /tag= af
FT FT /standard_name= "ORF L10"
FT FT /product= "protein of unknown function"
FT FT /note= "homologous to the NifX gene"
FT CDS 459579..460067
FT FT
FT FT /tag= ag
FT FT /standard_name= "ORF L11"
FT FT /product= "protein of unknown function"
FT CDS 460501..460920
FT FT
FT FT /tag= ah
FT FT /standard_name= "ORF L12"
FT FT /product= "protein similar to part of the Fe protein
FT FT of nitrogenase"
FT CDS 461228..461545
FT FT
FT FT /tag= ai
FT FT /standard_name= "ORF L13"
FT FT /product= "protein of unknown function"
FT CDS 463201..464739
FT FT
FT FT /tag= aj
FT FT /standard_name= "ORF L14"
FT FT /product= "peptidase-like protein"
FT FT /note= "homologous to the bi-Mpp gene"
FT CDS 464736..466079
FT FT
FT FT /tag= ak
FT FT /standard_name= "ORF L15"
FT FT /product= "processing protease-like protein"
```









-----

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 19:23:31 ; Search time 7811.87 Seconds  
(without alignments)  
4387.895 Million cell upd

Title: US-09-882-694A-2

Perfect score: 1638  
Sequence: 1 atgtctggccaccagcaactc.....tgcttcattgctcctgctaa 1638

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*

1: qb.ba:\*  
2: qb.btg:\*  
3: qb.in:\*  
4: qb.om:\*  
5: qb.ov:\*  
6: qb.pat:\*  
7: qb.ph:\*  
8: qb.pl:\*  
9: qb.pr:\*  
10: qb.ro:\*  
11: qb.st:\*  
12: qb.sy:\*  
13: qb.un:\*  
14: qb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*  
29: em.vi:\*  
30: em.htg.hum:\*  
31: em.htg.in:\*  
32: em.htg.other:\*\*  
33: em.htg.ov.inv:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	100	100	100	100	100	100

## ALIGNMENTS

RESULT	1	
AX076843		
LOCUS	AX076843	1638 bp
DEFINITION	Sequence 2 from Patent WO0105980.	DNA
ACCESSION	AX076843	
VERSION	AX076843.1	GI:13121518
KEYWORDS		
SOURCE	Exophiala spinifera.	
ORGANISM	Exophiala spinifera.	
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothryliomycetes;	
	Chaetothrichiales; Herpotrichiellaceae; mitosporic	
	Herpotrichiellaceae; Exophiala.	
	1 (bases 1 to 1638)	
REFERENCE	Duvick J.P., Maddox J., Gilliam, J., Folkerts, O. and Crasta, O.R.	
AUTHORS	Compositions and methods for fumonisin detoxification	
TITLE	Patent: WO 0105980-A 2 25-JAN-2001;	
JOURNAL	Pioneer Hi-Bred International Inc. (US); Curagen Corporation (US)	
FEATURES	Location/Qualifiers	

	1	1638	100.0	1638	6	AX076843	Sequence
	1	1638	100.0	1638	6	AX076843	AX076843
	2	1575	96.2	1691	6	AX076842	AX076842
	3	309.4	18.9	3715	1	CTE418060	Comamonas
	4	235.6	14.4	1842	1	ABD10439	Rhodococcus
	5	227.6	13.9	11471	1	AF257215	Brevibacterium
	6	210.2	12.8	11094	1	AE050584	Caulobacter
C	7	191	11.7	2618	1	RRS303350	Rhodococcus
	8	190.8	11.6	4170	1	RFL418061	Xanthobacter
	9	175.8	10.7	3584	1	RRH418062	Rhodococcus
C	10	169.8	10.4	12348	1	AE050586	Caulobacter
	11	158	9.6	2128	1	ACCCHM	M19029 Acinetobacter
	12	153.2	9.4	6330	1	AB060902	AB060902 Acinetobacter
	13	153.2	9.4	16160	1	AB026668	AB026668 Acinetobacter
	14	153.2	9.4	17417	1	AF282240	AF282240 Acinetobacter
	15	140.6	8.6	10549	1	AF2827214	AF2827214 Brevibacterium
	16	119	7.3	1642	6	423425	423425 Oligonucleotide
	17	119	7.3	1642	6	156573	156573 Sequence 2
C	18	90.8	5.3	16727	1	AE007015	AE007015 Mycobacterium
	19	90.8	5.3	39150	1	WFC212184	WFC212184 Mycobacterium
	20	86.8	5.3	10174	1	AE004637	AE004637 Pseudomonas
C	21	84.2	5.1	39446	1	SC87	AL132674 Streptococcus
C	22	78.2	4.8	13946	1	SC87	AE007131 Mycobacterium
C	23	78.2	4.8	16360	1	AE007131	AL132128 Mycobacterium
C	24	75.2	4.6	9776	1	MTV012	AB026669 Acinetobacter
	25	70.4	4.3	16527	1	AE004551	AE004551 Pseudomonas
	26	69.4	4.2	11354	1	AE005925	AE005925 Caulobacter
C	27	64.8	4.0	9902	1	SPP23219	AJ223219 Sphingomonas
C	28	64.2	3.9	10762	1	AE005810	AE005810 Caulobacter
C	29	63.8	3.9	11627	1	AE004582	AE004582 Pseudomonas
C	30	60	3.7	15124	1	AE006978	AE006978 Mycobacterium
	31	60	3.7	37630	1	MTX331	273101 Mycobacterium
	32	58	3.5	177466	1	AE001826	AE001826 Deinococcus
C	33	57.2	3.5	15019	1	AE006957	AE006957 Mycobacterium
C	34	57.2	3.5	29550	1	MTX039	AL132194 Mycobacterium
C	35	57	3.5	38209	1	SC5G9	AL117385 Streptococcus
C	36	54.8	3.3	3223	1	AE0030329	AF090329 Pseudomonas
C	37	54.6	3.3	110000	2	LMFLCRR18_00	AF099620 Leishmania
C	38	54.2	3.3	1323	1	AF355751	AF355751 Pseudomonas
C	39	54.2	3.3	14282	1	AE007188	AE007188 Mycobacterium
C	40	53.8	3.3	37751	1	MTX11A6	AF169016 Brevibacterium
C	41	52.8	3.2	15000	8	AF169016	AF169016 Aspergillus
C	42	52.6	3.2	339485	8	AF423282	AF172282 Oryza sativa
C	43	50.2	3.1	125020	9	AF2429315	AF2429315 Homo sapiens
C	44	50.2	3.1	339485	8	AF172282	AF172282 Oryza sativa
C	45	49.4	3.0	51300	1	AE241171	AE241171 Pseudomonas

DB	961	GATCTCATGGCTGCTCTGAGAGCCCGGTACTGGTTCGGTACCAAGCGCTCCCACTGGAG	1020
QY	1021	agcgaactactacgaatgctggagacggccggcgctcgaaattgtgaatctagaacaatcgg	1080
DB	1021	AGCGACTACTACGAATGCTGGANAGCCGACGCGTGGAAATGTGAATCTAGAACAATCG	1080
QY	1081	cccatttggcgtgtacaagaacagctgtgctctcttgagtgacggcgacgaagggaatgc	1140
DB	1081	CCCATTTGGCTGTTCAAAGACAGGTGTCTTGTGAGTGACGGCAGCAAGAGGGAATGC	1140
QY	1141	gacgcactgctgctggcgcggggttcgcagagctttcactgctcattgacacatatgggc	1200
DB	1141	GACACGATCTGTCTGGCGAGGGGTTTCGACAGCTTTCACCTGGCTCATTGCACACATATGGC	1200
QY	1201	ttgaaaaaacagcacggagtggacctgaaggagggtgtggaagaatggcatctacttat	1260
DB	1201	TTGAAAACAGCACGGGTGGACCTTGAAGAGGTGTGGAAGATGGCATATCTACTTAT	1260
QY	1261	atgggagctctctctcatggcttccccaatgctcttcttgctgcacggctcaagcccgcg	1320
DB	1261	ATGGGAGCTCTCTCATGGCTTCCCAATGCTTCTTCTGTCGGCACGGCTCAAGCCCGC	1320
QY	1321	accgctcttccaaacgcccacagcatcagaaaccccaagtcgacttgatcgcgataca	1380
DB	1321	ACCCTCTCTTCCAAACGCGCCACAGCATATAGAAACCCCAAGTCGACTTGATCGCGGATACA	1380
QY	1381	attgcaagtgtgagggccgagcacgccgctcgttgaggcgacgaatacagcaagaag	1440
DB	1381	ATTGCAAGTTGGAGGCCGAGCACGCCACGTCCTGTTGAGGCGACGAATACGACACAAG	1440
QY	1441	gcattgcgaattatattccaaagatgacagcacactcgttcccttgacggattcg	1500
DB	1441	GCATGTCTGATATATGATTGCCAAGATGAACGAGACACATCTGTTCCTTGCACGGATCG	1500
QY	1501	tggtagactggaggcaacatccctgggaagcaacacgtgtcttaacctcaatagcgcg	1560
DB	1501	TGGTGGACTGGAGGCACACATCCCTGGGAAGGCAACACGTGCTTTAACTTCATAGGCGGG	1560
QY	1561	attgctctctatgacgagatctgtcgaagaaaggctggccaattgggatgggtttgatgtg	1620
DB	1561	ATTGCTCTCTATGAGCAGATCTGTCAAGAGAAGGTGGCCAATGGGATGGGTTTGTATGTG	1620
QY	1621	cttcattgctccctgctaa	1638
DB	1621	CTTCATGCTCCCTGCTAA	1638

AX076842	1691 bp	DNA	linear	PAT 22-FEB-20001
LOCUS				
DEFINITION	Sequence 1 from Patent WO0105980.			
ACCESSION	AX076842			
VERSION	AX076842.1	GI:13121517		
KEYWORDS				
SOURCE	Exophiala spinifera.			
ORGANISM	Exophiala spinifera			
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothryiomycetes;			
AUTHORS	Chaetothryiales; Herpotrichiellaceae; mitosporic			
TITLE	Herpotrichiellaceae: Exophiala.			
JOURNAL	1 (bases 1 to 1691)			
FEATURES	Duvalick,J.P., Maddox,J., Gilliam,J., Folkerts,O. and Crasta,O.R.			
source	Compositions and methods for fumonisin detoxification			
	Patent: WO 0105980-A 1 25-JAN-2001;			
	Pioneer HI-Bred international inc. (US) ; Curagen Corporation (US)			
	Location/Qualifiers			
	1..1691			
	/organism="Exophiala spinifera"			
	/db_xref="taxon:91928"			
	/note="flavin monooxygenase with intron"			
BASE COUNT	397 a	444 c	480 g	370 t
ORIGIN				

L. 1691  
/organism="Exophiala spinifera"  
/db\_xref="taxon:91928"

Query Match		96.2%	Score 1575;	DB 6;	Length 1691;	
Best Local Similarity		96.9%	Pred. No. 0;			
Matches 1638;		Conservative	0;	Mismatches	0;	Indels 53; Gaps 1;
Qy	1	atgtcgccacacagcaactccagagcgaggttcttcgtcgatcgatcgagcgccatcatcgtt	60			
Db	1	ATGTGCGCCACACAGCAACTCCAGAGCGGATGTTCGTCGATCGATGCGAGCGCATCATCGTT	60			
Qy	61	ggagcgcgctcagcggcgcattctgtgtgtatcaaaatgcaaaagctcagactcaacgccc	120			
Db	61	GSAGCGCGCTCAGCGGCATCTGTGCTGTGTACAAATGCGAAAGCTCAGACTCAACGGCC	120			
Qy	121	aaaattctgagggagcccccgatttggcggcgctgtgacctggaaccgctaccctggc	180			
Db	121	AAAATCTTCAGAGGAGCCCCGATTTTGGCGCGCTGTGGCACTGGAACCGTACCCCTGGC	180			
Qy	181	gctcgtgttgatctggagacgcccctctaccactgaacattcccgaaagtatggaagac	240			
Db	181	GCTCGTGTGTATCGAGAGCGCCCTTCTACCAACTGAACATTCGCCGAAGTATGGAAGAC	240			
Qy	241	tgaacctgtcttgcgctatcctgaccagaaagagttgctgtcatatgttccacctgt	300			
Db	241	TGGACCTGTCTTGCCTGTATCTTGACCCAGAAAGAGTTGCTGTCTATATGTTCCAGCACTGT	300			
Qy	301	gacaagatccgggcttgaaaagacgtctacttcgagatcgaggtggttgatcgccg	360			
Db	301	GACAAGATCCGGGCTTGAGAAAGACGTCCTACTTCGGGASCTGAGGTGTTGATGCCCGG	360			
Qy	361	tatgcagagatctggccacctggactgtcaagacgctcggtcgccatgttgacagcga	420			
Db	361	TATGCCAGAGATCTGGGACCTGGACGTCTAAGACGCTCGGCTGTGGCATGTTCGACGSGCA	420			
Qy	421	aagtatctattctcgtacgggttgcctccacaggaagcacactcccgcaactcccgcc	480			
Db	421	AAGTATCTATCTCTACCGGGGTGCTCCACAGGAAGACACACTCCCGCACTCCCGCGC	480			
Qy	481	ctcgcgattcaacgggaagtgattcatctcgagtcgctggcgacgaagacttcgacga	540			
Db	481	CTCGCGATTTCAACGGGAAGGTGATTTCATTTGAGTTCGCTGGCGACGAAGACTTCGACGCA	540			
Qy	541	gagggccagagagtcgctcatcggtgcgcggggcccaagcatccagattgttcacggag	600			
Db	541	GAGGGCCAGAGATCGCGCGTATCGGTGCGCGGGGCCACAAGCATCCAGATTGTTCAGGAG	600			
Qy	601	ttagccaagaagctgaccaggttaacatgtttatgcaaaagccgagctattctgccc	660			
Db	601	TTGGCCCAAGAGCTGACCAAGTAACTTATGTTATGCGAAGCGCGAGCTATTGCTGTGCC	660			
Qy	661	atcgccgaacgaagatgtagaagaacgaagacgctggaaagccctactaccaccag	720			
Db	661	ATCGCGCAACGAACATGGATAGGAACGAACAGACAGCTGGAAGCGCTACTACCCGACG	720			
Qy	721	ctgtttgaagcagtcgaagctctcggttgattcccggtccagggcaccgctcggttggc	780			
Db	721	CTGTTTGAAGCAGTCGAAGCTCTCGGATTTGGATTCCCGGTCCAGGCACCGCTCGGTGGC	780			
Qy	781	atcttgaagtcagccccgagcagcgaggcgctatttcgaagagttgtggagcgtggg	840			
Db	781	ATCTTTGAAGTCAGCGCCCGAGCAGCGGAGGCGCTATTTCGAAGAGTTGTGGGAGCGTGGG	840			
Qy	841	gcctttaattttctgtgctgacgacgagagatcatgtttgacaaaagcccaaccca	900			
Db	841	GCCTTTAAATTTCTGTGCTGACGACGAGAGAGTCAATGTTGACAAAAGGCCAACCGA	900			
Qy	901	ctggtctatgactctgggccaacaaagactcgtctcgtatcgtcaatccgccaagaga	960			
Db	901	CTGGTCTATGACTCTTGGGCCAAAAGACTCGATCTCGTATGCTCAATCCGCCAAGAGA	960			
Qy	961	gactatggtcctctgagcccgctactggttcggtaccagcgtcccccactgag	1020			
Db	961	GATCTATGCTCTCTGAGCCCGCTACTGTTGCTGGTACCAACGCTCCCGCTCGGAG	1020			
Qy	1021	agcgactactacgaatgtgtggaacagcgagcgtcgaaattgtgaaatcttagaacaatcg	1080			
Db	1021	ACCGACTACTACGAATGCTGTGGACAGCGAGCGTCTCGAATTTGTGAATCTAGAACATCG	1080			
Qy	1081	ccatttggtggtgttacaagaacaggtgtccttgaagtcacgcaagaaggaatcg	1140			
Db	1081	CCATTGTGCTGTTTACAAAGACAGGTGTCTTTGAGTGCAGCAGCAAGAGGGAATCG	1140			
Qy	1141	gacaagatctgtcgtgcyagcgggttttcgacagdttcactggctc	1184			
Db	1141	GACACGATCGTGTGCGGACGGGTTTCGACAGTTTCACTGTGCTGTGAGTGTCTCGATC	1200			
Qy	1185	-----attgacacatatggggttga	1207			
Db	1201	ATGGTCTCGAGTCCGGACGTTTGGCTGACCTTTGAAAGATTGACACATATGGGCTTGAAA	1260			
Qy	1208	acaagcacgggtggaactgaagaggtgtggaagatggaacatactacttataatggag	1267			
Db	1261	ACAAGCACGGAGTGGACCTGAAGGAGGTGTGGAAGATGGCATATCTACTTATATGGAG	1320			
Qy	1268	tcttctctcattggcttcccccaatgctcttcgtcgccacgctcaagcccccgacctcc	1327			
Db	1321	TCTTCTCTCATGGCTTCCCAATGCTTCTGTCGCCACGCTCAAGCCCCGACCGTCC	1380			
Qy	1328	tttccacggcccaacatcatagaaccacgaatcgactgactgacgcgatacaatlgcaa	1387			
Db	1381	TTTCCAAAGCGGCCAACCATCATAGAAACCAAGTCACTTGCATGCCGATACATTTGCAA	1440			
Qy	1388	agttgagacgcagacacgcacgtccgttgagggcagcaaaatcagcaagaagaggtgt	1447			
Db	1441	AGTTGAGGCGGACGACGCCACGCTCGTTGAGGCGACCAATCAGACAGAGGCAATGGT	1500			
Qy	1448	cgattatgattgcaagatgaacgagcacactctgttcccccttgacggatctgtgtgga	1507			
Db	1501	CGATTATGATTGCCAAGATGAACGACACACTCTGTTCCTTGCAGGATTCGTGTGGGA	1560			
Qy	1508	ctggaggacaacatccctgggaagcaacgctgttacccttcatagggcggtatgcctc	1567			
Db	1561	CTGGAGGCAAGATCCCTGGGAAACACAGCTGCTTTAACCTTCATAGGGGGATTGCTC	1620			
Qy	1568	tctatgacagatctgtcaagagaaggtggccaattggagtggttgaagtcttcatg	1627			
Db	1621	TCTATGACAGATCTGTCAAGAGAAGTGGCCAAATTTGGATGGGTTTGATGTCTTCATG	1680			
Qy	1628	ctccctgctaa	1638			
Db	1681	CTCCCTGCTAA	1691			
RESULT 3						
CTE418060		3715 bp	DNA	linear	BCT 13-NOV-2001	
LOCUS		Comamonas testosteroni cpmA gene for cyclopentanone				
DEFINITION		1,2-monooxygenase and cpMB gene for cyclohexanol dehydrogenase.				
ACCESSION		AJ418060				
VERSION		AJ418060.1				
KEYWORDS		cpmA gene; cpMB gene; cyclohexanol dehydrogenase; cyclopentanone				
SOURCE		1,2-monooxygenase.				
ORGANISM		Comamonas testosteroni.				
REFERENCE		Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;				
AUTHORS		Comamonas.				
TITLE		1 (sites)				
JOURNAL		van Beilen, J.B., Fritzsche, U., Seeger, M., Smits, T.H.M. and				
AUTHORS		Witholt, B.				
TITLE		Cloning of Baeyer-Villiger monooxygenases from Comamonas,				
JOURNAL		Xanthobacter and Rhodococcus via PCR with highly degenerate primers				
AUTHORS		unpublished				
TITLE		2 (bases 1 to 3715)				
JOURNAL		van Beilen, J.B.				
AUTHORS		Direct Submission				
TITLE		Submitted (01-NOV-2001) van Beilen J.B., Institute of				
JOURNAL		Biotechnology, ETH, ETH-Hoenggerberg, CH-8093, SWITZERLAND				
FEATURES		Location/Qualifiers				

source	1. .3715 /organism="Comamonas testosteroni" /isolate="NCIMB 9872" /db_xref="taxon:285" /note="this strain was previously named Pseudomonas sp." NCIMB 9872" 502. .2154 /gene="cpmA" 502. .2154 /gene="cpmA" /gc_number="1.14.13.22" /function="converts cyclopentanone to valerolactone" /codon_start=1 /transl_table=11 /evidence=experimental /product="cyclopentanone 1,2-monooxygenase" /protein_id="CAD10799.1" /db_xref="GI:16943679" /translation="MTTMTTTEQLGMNNSVNDKLDVLLIGAGFTGLYOLYHLRLKLG YKHLVDAGADIGGIHIMNCYPGARVDTHCQYQYSIPELWQEFNMKELFPNNAQMR YFHFADKLDLSKDISFNTRVOSAVDFEGTREMVTVRISGHOPQIARFVIANLFGCAS STPNVDGIETFKQWHTALMPOQGVNMGAKRVALIGTSSGVQVQVADALDAKQVTV YORTPNLALPHQKQSLAEDNLRMKPELPAAFERRGCKFAFGDFDTAKNATELSAAM RTEILEELMNAAGSPRYLANFDQYLLDDKANDYVYEFWRQKVRARIKDKVAELKAPM KHPHYGAKRRPSLEQWYIEIFNONNVTLVDVNETVPLRITKEGIVTAEGEAFDLIVF ATGPDATGGTSDIFDRNQGSFQVMSDGIQTLQVATAGFPNLIFGYGPOSFAGF CNGSSAIEQCDLILQIMNYLRDNNISRTAQSEAEQWENKLIADFWDSLSLPPRAKSW YQGSNIPKVKYSELNPPGLPTTYISKPNESAEGKVGAGFSLAS" 2187. .2939 /gene="cpmB" 2187. .2939 /gene="cpmB" /gc_number="1.1.1.1" /function="oxidation of cyclic alcohols to the corresponding ketones /note="similar to short-chain alcohol dehydrogenase family" /codon_start=1 /transl_table=11 /evidence=experimental /product="cyclohexanol dehydrogenase" /protein_id="CAD10799.1" /db_xref="GI:16943680" /translation="MGRVNDKVLVLTGGAMGMLTHCTLLAREGATVYLSDMNEELGH QAVAEIRROGSKAHFLIDVTNENHMTGAVDTILAESDRDLALVNAGLITLKPQVQDT SNEWDRIFETINRSVFLGTRAVIPMRKAHKGCIVNVSYIQLVGPAGAAAYEASKG AVRLFTKACAVDLAPFNIRVNSVHPGVATPMTQIILDAPOSARALLGPTLLGRAAQ MEVSOAVLFLVSDSEASVHGSSELVVDGGYTAN" 2 others BASE COUNT 823 a 1113 c 1030 g 747 t ORIGIN
gene	808 TATTTCCATTTTGGCCGACAAGAGCTCGAGCTGAGCAAGGACATCAGCTTCAACACCGT 867 346 GTGGTGTGATGCGGCGGATGATCCAGAGAGTCTGAGGACCTGAGGAGCGTGGTGGTGGT 405 868 GTGCAAGTCGGCGGCTCTTTGAGCAAGGACACGCGGAATGAGCGGTAGCTCGATCGACAC 927 406 CATGTTGCGAGCGCAAGTATCT 465 928 CAGCCGATCCAGCCAGGTTGCTCATGCCAACCTTGGCTTCGGTGCACGACCCAGCAGC 987 466 CCGGCACTCCCGGAGCTCCGAGATTCAGCGGAGAGTgattcattcagagtgctcgg--- 522 988 CCCAATGTGATGGCATCGACATTTCAAGGCCGATGCTGATCACAACCTCTCTGTGGGCC 1047 523 CAGCAAGATTCTGAGCGAGGAGGCGGAGAGTGCCTGATCGTCTCTCTCTCTCTCTCTCTCT 582 1048 CAGCAAGGCGTGAACATGCGCGGCAAGCGGCTGGCCATCATTTGGCACCGGCTCCAGCGG 1107 583 ATCCAGATTGTCAGGAGTGTGGCAAGAGGCTGACCAAGTAACTGTTTATGCGAAG 642 1108 GTCCAGGTGCGCCAGGAGGCTGCTTTCATCCGAAACAGGTGACGCTGTACCCAGCGCAC 1167 643 CCGAGCTATTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702 1168 CCCNACTGCGCTTGGCCATGCATCAGAACGAGCTCAGCGCGGAGGACAAATCTCGGCATG 1227 703 AAGGCTCTACTACCCAGCTGTTTGAAGCGAGTGAAGTCTCGGATGAGTCTCTCTCTCTCTCT 762 1228 AAGCGGAGCTTCCGCGACGCTTCGAGAGAGCGCGGCAAGTGTCTCCCGGCTTCGACTTC 1287 763 CAGGCACTCGTGGTGGATCTTGAAGTCAAGCCCGAGCGCGGAGGAGGAGGAGGAGGAGGAG 822 1288 GACTTCATCCCAAGAACGCGAGCTGCTGCTGCGGAGCGGACAGAGATCTTGGAA 1347 823 GAGTGTGAGAGCGTGGGCGGCTTCCGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882 1348 GAGCTGTGAAGCGCGCGGCTTCCGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1407 883 GACAAAGGCGCAAGCGAGTGTGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 942 1408 GATGACAAAGCGCAAGGATTTAGCTGCTAGAGTGTGCTGCGGCGGACAAAGTCTCGGCG 1467 943 GTCAATCCCGCAAGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1002 1468 AAGGATCCGAAAGTTGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1527 1003 AAGCGCTCCCGAGAGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1062 1528 AAGCGCGCTTCCGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1587 1063 GTGAATCTAGAACAATCGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1122 1588 GTGGATGTCAAGCAAGCAAGCGGCTTCCGATCAGCGAGAAAGG---CATGTCAGCGCT 1644 1123 GCGAGCAGAGAGGAGTGTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1182 1645 GAGGTTGAAGCGCAAGTTCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704 1183 TCAATGACACATATGGCTGAGAAACAGCAGCGAGTGTGAGCTGAGGAGGAGTGTGAGAA 1242 1705 GGACTCACCAGCATGCTGCTGCGCAACAGCAGGCGCAGAGGCTTCAAGGATGTGTGCTCT 1764 1243 GATGGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302 1765 GACGGAATCCGCAAGCTGCTGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824 1303 GCCAGGCTCAGCGCGAGCGCTCTTCCAGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1362 1825 TAGGGACCTCAATCGCTGCGGCTTCTGCAACGCGTCTGCGAGCGCGGAGTACACAGGCG 1884 1363 GAGTGTGAGCGCGAGTCAATATGCAAGTGTGAGGCGGAGCGAGCGAGCGAGCGAGCGAGCG 1422
gene	106 ctagactcaacgcaaatcttcgagggagcccgatttggcgcgctctggcactgg 165 628 CTGGGCTACAAAGGTTTCATCTCGTGCAGCGCGGTCGCGGATATTGGCGGGATCTGGCA 687 166 aacgcgtacctggcgctgtgtgattgagacgacgaccttctaccgaactgaacatctccc 225 688 AACTGCTACCGCGAGCGGTGGATACCCACTGCGCAGATGTACACAGTACTCCATCTCCA 747 226 gaagtatgaagactgacatgcttgcgctatctcagcagcaagagattcgtca 285 748 GAGTGTGCGCAGGACTTCAACTGGAAGAGCTGTTCCCTAACTAGGCGCAATTCGCGAG 807 286 tatgttcaccactgagcaagactccgggcttgagaaaagactctactctcggagctgag 345
CDS	Query Match 18.98; Score 309.4; DB 1; Length 3715; Best Local Similarity 51.48; Pred. NO.1.5e-68; Matches 765; Conservative 0; Mismatches 716; Indels 6; Gaps 2; Qy 46 gacgcacatcgttggagcgcgctcagcgcatctctcgtgtatcaaatctgcgaaag 105 Db 568 GACGCTTTCATCGCGCGCGGCTTCCCGGCTCTTACAGCTCTATCAGCTCGGCAAG 627 Qy 106 ctagactcaacgcaaatcttcgagggagcccgatttggcgcgctctggcactgg 165 Db 628 CTGGGCTACAAAGGTTTCATCTCGTGCAGCGCGGTCGCGGATATTGGCGGGATCTGGCA 687 Qy 166 aacgcgtacctggcgctgtgtgattgagacgacgaccttctaccgaactgaacatctccc 225 Db 688 AACTGCTACCGCGAGCGGTGGATACCCACTGCGCAGATGTACACAGTACTCCATCTCCA 747 Qy 226 gaagtatgaagactgacatgcttgcgctatctcagcagcaagagattcgtca 285 Db 748 GAGTGTGCGCAGGACTTCAACTGGAAGAGCTGTTCCCTAACTAGGCGCAATTCGCGAG 807 Qy 286 tatgttcaccactgagcaagactccgggcttgagaaaagactctactctcggagctgag 345







```

/codon_start=1
/transl_table=11
/product="putative regulator"
/protein_id="AK73168.1"
/db_xref="GI:14719382"
/translation="MRTLDPHSRDEMRLDTVLAAALADPVRRTVACKLNDAPGDHACAP
FELPVSKEATATYHFTLRAGVIRQYEGTKIMNTCARTIDARPPGSRVFAQDIER
AEAAAPESH"
8198. .9271
CDS
/note="similar to putative flavin oxidoreductase proteins"
/codon_start=1
/transl_table=11
/product="NADH-dependent flavin oxidoreductase"
/protein_id="AK73169.1"
/db_xref="GI:14719383"
/translation="MSHLLEPELTIRGLTFRNRHVPVPCQYSVETLDGVPAPHVTH
YGAARRGAGAVEATGVAPEARISAKDLGNDEQRDAFPIVDVLTGGAAGIQL
AHARRASTYEPWGTDRGSLPVDGEGWTAPSAALFDGLAEPRALTEIIEVVA
FRSSARRATEAGDFVEIHAAGHLLHEFLSPLNNRTDSYGGSLNKRARLLIDIVA
TRTEVEDVPVFLRSATDWTGGGLTLDITVEVAGWLKEHGVLDLIDVSSGQNVWASP
VGPYOTLAAAGVROGSGLPTAAVGLISEPFGHEHILATGOADVLLVGREYLRDPEA
LRAADALRFDIDYRPAQYHRAVK"
complement(9363..10154)
CDS
/note="similar to gntR-family transcriptional regulators"
/codon_start=1
/transl_table=11
/product="putative regulator"
/protein_id="AK73170.1"
/db_xref="GI:14719384"
/translation="MCRGSSRSTSGSPSCACWIPHRAAPMKGVARGSKSADP
SWADSHHTANWPIRRAASARVRVGRAPASGLPRNARYSSAVSPSPSPRRRCV
EEVDDAAGLRCTCTRRRTTIRLITNADRTHOVPEHGPVAGNDHHCGLGTRAPAD
RTGRSSASLAPGARHGIPLRLYDVFDEGLRNGPASPFTSLGRMRQGRARHPRSH
DGRDNVARRPGLAATDRSELSSHSTRSPSASAVRADRRQPS"
complement(9832..10437)
CDS
/note="similar to CrcB"
/codon_start=1
/transl_except="(pos:complement(10435..10437),aa:Met)
/transl_table=11
/product="putative CrcB-like protein"
/protein_id="AK73171.1"
/db_xref="GI:14719385"
/translation="MLQWGRDPTVKAPSPRRRSAGDRHRRHGPRIPODVADPNR
FPCRRSSRRGRVTVPVHLRLSYLGIAF IGGAAGTAAREASISLPLGDPWPAILT
VNILGAFLLGLDLSIARSGPDRWRRTIRILVGTGPMGFTTYSALAAADTAGRLGAG
PGASSPGIGIAVGTGVLIGLATFAGIATATTLRRGGAR"
BASE COUNT 2035 a 3683 c 3786 g 1967 t
ORIGIN

Query Match 13.9%; Score 227.6; DB 1; Length 11471;
Best Local Similarity 48.4%; Pred. No. 1.6e-47;
Matches 725; Conservative 0; Mismatches 764; Indels 9; Gaps 3;

QY 35 ccgtcgatcgagccatcatctgttagcgccgctcagcgcatctctgtgtgaca 94
DB 5891 CCGAGGTCCTGACGCATCATCTGGTGGCGGATTCGGGGCCCTGTATCTGTCGACC 5950

QY 95 aattcgaaactcagactaacccaaatctctcagaggagcccccgatatttgcggcg 154
DB 5951 GCCTCGTGAAGACGGGTCAAGGTCTGGGACGCCGCCGGGATTTCGGGGCA 6010

QY 155 tctggcactgaacgcctaccctgcgctcgtgtgtattcggagagccctctaccac 214
DB 6011 TCTGGTGTGGAACCTGCTACCCGGGTCTCGTAGGACAGACACCGGACAGATCTATCAGT 6070

QY 215 tgaacatcccgagtagtgaagactggacctggtcttgcgctatctgaccagagaag 274
DB 6071 TCCAGTACAGAGACTCTGGGAAGACTTCGACTTCAGAGAGCTCTACCCCGACTTCAACG 6130

QY 275 agttgctgtcatatttaccactgtgacaagatccggggcttgagaaaagacgttact 334
DB 6131 GGGTTCGGGAGTACTTCGAGTACCTGCAGCTCGCAGCTCGACCTGTCCCGCGAGCTC 6190

```

```

QY 335 tcggagctaagtgattgatgcgagtagtccagagatctgggacacctggactgtcaaga 394
DB 6191 TCAACACCTTTGCGGAGTCTCTGCACATGAGACGCTGCCAAGAGATGGAGGTGCGAT 6250

QY 395 cgtcgctgccaatgtttgcagcggcaaatgtctctctctgctacggggttgcacaca 454
DB 6251 CGTCGAAAGACGTGAGCAGCGGCCCGCTGCGCTCATCTGTCGCCACCGGCTTCGGTGGCA 6310

QY 455 ggaagcacactcccgcaactcccgccctgcgcgatttcaacqgggaagtgattcattcga 514
DB 6311 AGCCCTCTACCGAATCGAGGGCTTCGACAGCTTCGAAGCGGAGTCCACACACCG 6370

QY 515 gtccctg--gacgaagacttcgacgaagagggccagagagatcgccctcatcggtccg 571
DB 6371 CAGCGTGGCGCAGGCTGCGCTCGACATGACGGGCAAGCGAGTCTGCTCATGAGCGACCG 6430

QY 572 gggccacaagcatccagattgttccagagattggccaaagaagtgaccaggttaaccatgt 631
DB 6431 GTGCTTCGGCATCCAGGTCAATTCAGAAAGCGCGCGGTTGCGGAACACCTCACCGTCT 6490

QY 632 ttatcgaaagggcgagctattgtctgcccactgcgcaacgaacgatgataaggaacaac 691
DB 6491 TCCAGCGCACCCCGAACCTTGGCCTGCCATGCGGACGAGCGGCTGTGCGCCGATGACA 6550

QY 692 agacagcctggaagcctactacccacgcgtgtttgaagcagtcgaaagtctcgattg 751
DB 6551 ACGATCGTACCAGAGAAACATCGAAGATCTGTTTCCAAATCGGTGACAAATTCGTTTGGCG 6610

QY 752 gattccgggtccagcagccgctggtggcatcttgaagtcagccccgagcagcgggag 811
DB 6611 GATTGCACTTCTTCTATCCCGACAAAGCGCGGACACACCCCGGAGACGACGCGGACCG 6670

QY 812 cctatttcgaagagtggtgggagcgtggggcccttaatttcttctgttgcagttaccgag 871
DB 6671 CGATCTACGAAAGATGTGGGCAAGGCGGATTCCTACACTGTGTGGTTCGGAACACTTCCAGG 6730

QY 872 aagtcagtggtgacaaaagggccaaacgactggtctatgactcttggggccaaaagactc 931
DB 6731 GACTCTCACCGATGAGGACGACCAACACACCTTCTTACAACTTCTGGCTTCGAAGGTGC 6790

QY 932 gatctcgatctgtaactccggcaagagagatctctatggctctctctgagcgcgcgtact 991
DB 6791 ACGATCTGTGAAGGATTCGCAAGACCGCGGAGATGCTCGACCGGCGACCCCGACCGACC 6850

QY 992 ggttcgtatcaacagcgtcccccactggagcagactactacgaaatgctggacaagcga 1051
DB 6851 CGTTCGGCGTCAAGCGTCCCTCGCTCGAACAGAACTACTTTCGACGTATATACACAGGACA 6910

QY 1052 gcgtgaaattgtgaatctagaacaatcgccatctgggctgttacaagacaggtgtgc 1111
DB 6911 ATGTGATCTCATCTGACTCGAATGCGCACCCCGATACCCCGGTCTCTTCGAACGCGGTGC 6970

QY 1112 tcttgatggcgcgacgaagaggaatcgacacagctgctgtgctgcagcgggtttcgaca 1171
DB 6971 AAACCCCGGACGGAGTGTGTC--GAATGCGATGCTCTGCTGTGGCCACCGGCTTCGACA 7027

QY 1172 gtttctactgctcatgacacatatgggttgaaacaaagcagagtgagcctcgaag 1231
DB 7028 ACAACACGGCGGATCAACGCCATCGATATCAAGCC--GGCGGCGAGCTGCTGCGGTG 7084

QY 1232 aggtgtgaaagatggcatatctatattatggagctcttctctctatggcttcccccaatg 1291
DB 7085 ACAAGTGGCGACCGGCTGGACACCTACATATGCGGCTGTGCGACGCGGAGGATTCGCCAATC 7144

QY 1292 cttcttcgtcgcaagcgtcgaagccccgacgctcttccaaacgcccccaacgacatag 1351
DB 7145 TCATGTTCTCTACGGCCCCGACAGCCCTTCGGGCTTCTGCAATGGGACCGGACTTCGGCG 7204

QY 1352 aacccccgaagctgactgacgcgacatacaattgtcaagttggaggccgagcagccagct 1411
DB 7205 GAGCGCCAGGCGATATGTTGCGCGGACTTCTCTATCTGCTGCTCAAGGACACACGSCATCTCGC 7264

```







[illegible]

/product="cyclohexanol dehydrogenase"

/protein\_id="CAD10802.1"

/db\_xref="gi:18495819"

/translation="MKRVQKVALVTGGAMGRTHAETLAAEGAYVFFVGRDRAAAGK

AVADGIVAKGKABFLSDVTKESDAAAVDTKAKSGKLNVLNAGTLVLKPAHET

TNDWDSTFNVNRYGVLGIRSFVPLMKNGNSIINISITYGIVGAPMAGAVIASKGA

VRLLFKSCAVDLIOFGIRVNSVHPGVIDTPTMTDLLHGPAPARKAIMGATLDFRCQP

VEVSKAVFLASDESSRFVHGAEWYVD\*

BASE COUNT 857 a 1247 c 1331 g 735 t

ORIGIN

Query Match 11.6% Score 190.8; DB 1; Length 4170;

Best Local Similarity 47.5%; Pred. No. 4.2e-38;

Matches 712; Conservative 0; Mismatches 767; Indels 21; Gaps 4;

Qy 36 cgtgcacgacgcacatcatctgttgagccgagcctcagcgccatctctgtgtacaa 95

Db 1760 CGCGGATTATGACCGCGTGTCTCGGTGCAGGGTTCGGCGGACTCTATCGCGTGACAA 1819

Qy 96 attgcgaag---ctcagactcaagcccaaatcttcgagggagcccccattttgacg 152

Db 1820 GTTGGCCATGAACAGGGGATGAACGTCAAGGCTATGACAAATGCCGACACATTCGGCGG 1879

Qy 153 cgtctggcactgaacccgtacccctggcctgctgtgtgattcggagacgccccttcacca 212

Db 1880 AACCTGGTTCGTGAACCGGTATCCGGCGCGGTCTCGGACACCCGAAAGCTTGTCTACCG 1939

Qy 213 actgaacattccc---gaagatggaagacgtggacctggtcttgcgcgtatcctcgaaca 269

Db 1940 CTTTTCCTTCGACAGGGAATTTGTCAGCGCGCGCTGTGAAGAACCGCTACGTGACCCA 1999

Qy 270 gaaagattgctgtatatttcaccactgtgacaagatccgggtgttgaaagacgt 329

Db 2000 GCGGAAATCTGGCTATCTCAAGAGGTGCGCGATCACTCAGCTCGCGCGACCTA 2059

Qy 330 ctacttcgaagctgaagtggttgatgctgcggtatgcccagagatctgggacacctggactgt 389

Db 2060 TGAATTTCAACACCAAGGTTCTCGCGCGCGAGTTTCGATGACGCAACCGCGCTGTGGAAGT 2119

Qy 390 caagacgtcggtcgccatgttgacgcgcaaaagtaactcattctcgtacgcggtgtgct 449

Db 2120 CACGACGACAGGAGACGGCGTTCACGCGCAAGTATCTGATACACCGCGCTAGGGCTTCT 2179

Qy 450 ccacggaagacacactcccgactcccgctcgcgatttcaacgggaagtgattca 509

Db 2180 GTCGGCACCTAACCTGCGGAAAGTTCAAGGGATGACACCTTCAAGGCGCGATCTGCA 2239

Qy 510 ttcgagtgctgcacgaagacttcgacgagagggccagagagtcgcctcactcgtgc 569

Db 2240 TACCGCGCTGCGCGAGGTGTGAACCTTCCCGCAAGCGTGTGCGCATCATCGGAAC 2299

Qy 570 cggggccacaagcatccagattgttcaggagttggccaaagggctgaccaggttaacct 629

Db 2300 CGGCTCAGCGGGTTCAGGTGATCACCGCCACCGCGCGGATTTGCCAAGCACTTGACCGT 2359

Qy 630 gttatgcgaagcgcagattgtctgcctcctcgcgcaacgacagatgatgaagca 689

Db 2360 GTTCCAGCGTTCGGACAAATTTGCTGTTCCATCGGCAACACGCGCAGGATCCGCAAC 2419

Qy 690 acagacagcctgaagggcactaccacgctgtttgagcgagtcgaagtcctggat 749

Db 2420 GATCGGAGCAGAGGCGACGTATGACGACATCTGGAACACAGGTAAATCTCGGCCCT 2479

Qy 750 tggattcccggtccagggcacctcgtgttggtgcatctttgaagtcagcccagcagcgga 809

Db 2480 GGCATTCGCGTTCGAGGGAAGCACTATTTCGGCGCGAGACCGCTCGCGCCGAGGAACGCA 2539

Qy 810 ggcattattcgaagagttgtggagcgtggggcccttcaattttcttgcctgcccagtaacg 869

Db 2540 CCGGATTCGAGGCGGCTGCGACGCGCGCGGCTTCTACTTCATGTTCCGAACGTT 2599

Qy 870 agaagtcaggttgacaaaagccaaaccgactggtctatgacttcttggtggccaaaagac 929

Db 2600 CTCGACATCGCTACAGTCAAGTCCGCAACGATCGCGTCCGCACTTTCATCAACGGGAA 2659

Qy 930 tgcgactcgtatcgtcaaatccggcgaagagagatctcattggtcctctctggagcgcgcta 989

Db 2660 GCTCAACGACATCGTGAAGGATCCCGA-----AACCGCGGGAAGCTGACCCCTC 2710

Qy 990 ctggttcggtaacaaagcgtcccccactgagagcgaactactacgaaatgctggacaagcc 1049

Db 2711 GGACCTCTATGCCAAGCGTCCGCTGTGCGGAGGACGACTATTATGTGCTCTACACCGGGA 2770

Qy 1050 gagcctcgaaattgtaatttagaacaatcgccatttggtgctgttcaaaagacaggtgt 1109

Db 2771 CAACGTCACGCTGGCGGATGTGAAGCCGACCGCATCGCGGAATTCACCCCTGACCGGCAT 2830

Qy 1110 gctcttgagtcagcggcagcaaggggaatgcgacacgactcgtgctggcgagcgggttcga 1169

Db 2831 CCGGCTTGGCAGCGCGCGCAACACGAACTGGACGTTGTGTGATCTTCCGCCACCGGATTCGA 2890

Qy 1170 cagttcactggtcattgacacatatggttgaaacaaacagcagcggagggacctgaa 1229

Db 2891 TGGCGTGGACGGCAACTACACCGCATGACATGCGCGCAGGACGCGTCACGCTGCG 2950

Qy 1230 ggaggtgtgaaagatggtcatttacttatatggtggttctctctctatggtctcccaaa 1289

Db 2951 CGACATGTGAAGGAGGGGCCCTTGGCTATCTGGGCATCATGAGGCGGAATTTCCCAAA 3010

Qy 1290 tgccttctcgtgcacggcgtcgaagccccgacctcctcttcccaagcggcccaacgactat 1349

Db 3011 CTTCTTCATGATCTTGGCCCGAAGCGCC-----CTTACCAACCTTCTCCACGAT 3084

Qy 1350 agaaacaaagtcactgtatgcgacataaatgcaagttggagggcgagcagccac 1409

Db 3065 CGACACGACGCTGGATGCGCATGATGCTCAAGACGATGGAAGCCAAAGGGCTGAA 3124

Qy 1410 gtccgtgagggcagaaatcagacaaagagcagtggtcgattgattggtccaaagatgaa 1469

Db 3125 GACCAGCAACCAACCGCGCAAGCTCGCATCAATGGTGGAACTGTGCCACCACTATCGC 3184

Qy 1470 cagacacactctgtccctctgacggttcgtggtggtgaggaacacatccctgggaa 1529

Db 3185 AAACATGACCTTGTCCGAGGAGAAAGCTGGATCTTCGAGCAATATACCGGCA 3244

# RESULT 9

RRH418062/c  
LOCUS Rhodococcus rhodochrous ORF1 for putative sterold monooxygenase /  
DEFINITION esterase fusion protein.

ACCESSION AJ418062

VERSION AJ418062.1 GI:18495820

KEYWORDS esterase fusion protein; ORF1; sterold monooxygenase.

SOURCE Rhodococcus rhodochrous.

ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Rhodococcus.

REFERENCE 1 (sites)

AUTHORS van Beilen, J.B., Fritzsche, U., Seeger, M., Smits, T.H.M. and Witholt, B.

TITLE Cloning of Baeyer-Villiger monooxygenases from Comamonas.

JOURNAL Xanthobacter and Rhodococcus via PCR with highly degenerate primers

REFERENCE 2 (bases 1 to 3584)

AUTHORS van Beilen, J.B.

TITLE Direct Submission

JOURNAL Submitted (01-NOV-2001) van Beilen J.B., Institute of Biotechnology, ETH, ETH-Hoenggerberg, CH-8093, SWITZERLAND

FEATURES

source

1..3584

/organism="Rhodococcus rhodochrous"

/isolate="DSM 11097"

/db\_xref="taxon:1829"

complement(160..2712)

CDS

```

/note="ORF1"
/codon_start=1
/transl_table=11
/product="putative steroid monooxygenase / esterase fusion
protein"

```

BASE COUNT	598 a	1163 c	1172 g	651 t
ORIGIN	QWRYRVRGAGARACGGAGNOVNRNDGRATSGATFETSLAQDVAIVGVIRAEVGNK			
Query Match	10.7%; Score 175.8; DB 1; Length 3584;			

Qy	46	gacgcacacacgtgtgagacgcgcgcctcagcggcactctctgctgtgtacaaaattcgcaa--	103
Db	2697	gatgtagtgatgtctggcagcgcgggttgcgcggctgttacgcgttccacaaagctgcgtaac	2638
Qy	104	-agctcagactcaacgcacaaatcttcagggaggccccgcattttggcggcgtctggcac	162
Db	2637	gaaactcgtcttggaagctgcagagcgtgtccaccccgccggcgccggcggaacgtgtgtgg	2578
Qy	163	tggaaacgctaccctggcgtcgtgtgtatttcgagacgccttctaccacactaaacatt	232
Db	2577	tggaaacgctattccggggcgacgcgtccgattttgcacttcgctctactctactctgcgttc	2518
Qy	223	cccgaag--tatggaaagactggacctggctctgcgcctatctctgaccacagaagattg	279
Db	2517	tccggaggagatccagccgcgagtgccagtgccagggaagtacccgcacacacgcgggatt	2458
Qy	280	ctgtcatatgttaccacactgtgacaagatcggggcgttgagaaaaagacactacttcgga	339
Db	2457	ctgaacttacctctgagtgccgtggcgacccgtcttcgacctctgaagccgtctccaatttcac	2398
Qy	340	gctgagtggttgatgcgcgggtatgccagagatctggcagccttggacttgaagactcgc	399
Db	2397	actcgggtgacctcgatgtgtcgaaacagacagaggcgacgacttggaccctctggcacccgac	2338
Qy	400	gctgggcattgttgacgcgcacaaagtattcattctcgtacgcgggttgctctccacggaag	459
Db	2337	gacggacgcagctgcacttcgcccgttacctgtctctcggaacacgcgcacactatctgac	2278
Qy	460	cacactccgcgaactccccggctcgcgcatattcaacggggaaggtattcattcctgagtgcc	519
Db	2277	aagaagccggattccccgggattcgaaagacttcaaggctgaggttctgcagacgacgactgc	2218
Qy	520	tg---gcacgaagacttcacgcagagagggcagagatccctcactcgttccggagcc	576
Db	2217	tggccgcacgacccgctgcattctaccggcagcgcgtctgggatctatcggaacccggtgcc	2158
Qy	577	acaagcatccagatgttctcagggttggccaaagaagcctgaccaggttaaccatgtttatg	636
Db	2157	accggaatccagactcattcccggaacttgcgaagcacgcgcgcacactctgacgtcttccag	2098
Qy	637	cgaaggccgcagctattcttgcccatctgcggcaacgaacgatgataggaacgaacagaca	696
Db	2097	cgcacccgcagcttcgcatccgcctgtgggaaacacgaaggttcgattcccgagcgcgcgca	2038











Db 859 ACCCGGTTCAATCGCGCTCATTAACACGAAGCAGATGCCCTGTGGGAAGTCACACACGAA 918  
Qy 400 gctggccatgttgcacgcaaaagtatctctattcctgcgtacaggggttgcctccacaggaag 459  
Db 919 TATGGTGAATAACACAGCGCGTGTCTCTCATCACTGCTTAGGCTTATGTCTGCCGCT 978  
Qy 460 cacactcccgccactcccgccgctgcgcgatttcaacggaaggtgattcattccatgcagtgcc 519  
Db 979 AACTTGCCAAACATCAAGGCAATTAATCAGTTTAAAGGTGAGTCATCATACACGCGC 1038  
Qy 520 tggcacgaagacttcgacgagagggccagagagtcgcctcatcgtgcgcgagggccaca 579  
Db 1039 TGGCCAGATGACGTAAAGTTTGAAGGTAAACGTGTGCGGCTGATTTGGTACGGGTTCCACC 1098  
Qy 580 agcatccagatgttcaagagttggccaagaagctgacacaggttaacacatttattacga 639  
Db 1099 GTGTGTCAGGTATTACAGCTGTGGCACCTCTGCTTAACACCTCACTGCTCTCCAGCGT 1158  
Qy 640 aggcagagctatgtctgcccactgcggcgaacacgatgtagaagcaaacagacagcc 699  
Db 1159 TCTGCACAATACAGCGTTCCAAATTTGGCAATGATCCACTGCTGGAAGAAGATGTTAAAG 1218  
Qy 700 tggaaagcctactaccacacgctgtttgaagcagagtcgaaagctctcggattggattcccg 759  
Db 1219 ATCAAGAGACAATATGACAAAATTTGGGATGGTGTATGGAATTCAGCCCTTGCCCTTTGGC 1278  
Qy 760 gtcaggcaccctcggttggcatctttgaagtcagcccgagcagcgagggagccctatttc 819  
Db 1279 CTGAATGAAGCACAGTCCAGCAATGAGCGTATCAGCTGAAGACGCAAGCAGGATTTT 1338  
Qy 820 gaagagttgtggagcgttggagcctttaaatttcttctgctgcagtcagcagaagtcag 879  
Db 1339 GAAAGCATGCAACAGGTGCGGTTTCGCTTCATGTTGAACTTTTCGGTGATATT 1398  
Qy 880 gtigacaaaagccaacgactgctatgaacttcgggcccgaagaaagactcgtatcgt 939  
Db 1399 GCCACCAATATGAAGCCAAATATCGAAGCCCAAAATTCATTAAAGGTAAATTCGTGAA 1458  
Qy 940 atcgtcaatccggcaagagagatctcatggtctctcttgagccgcgtgaggttcggt 999  
Db 1459 ATCGTCAAAAGATCCAGC-----CATTTGCACAGAAAGCTTATGCCACAGGATTTGAT 1509  
Qy 1000 accaagcgtccccactggagagcagactactacgaatgctggcaagcccgagcgtcgaa 1059  
Db 1510 GCAAAACGTCGTTGTGCAGAGTGTGTACTACAACACCTTTAAACCGTGAACAATGTCCT 1569  
Qy 1060 attgtgaatcagaacaaagcccatctgtgctgtttacaagacagaggtgtgctcttgagt 1119  
Db 1570 TTGAAGATGTGAAAGCCAATCCGATTGTTGAAATTAACGAAACGGTGTGAACATCGAA 1629  
Qy 1120 gacgagcagcaagaggaatgcacacatcgtgctgagcagcggtttcgacagattcact 1179  
Db 1630 AATGGCCATTTTCGTTGAATTAGACATGCTGATATGTGCCACAGGTTTGTATGCCCTCGAT 1689  
Qy 1180 ggtcattgacacataatgggcttgaacaaacagcagagtgaggacctgaagaggttgg 1239  
Db 1690 GGCACATATGTCGGCATGGACATTCAAGGTAAACAGCGCTTGCCATGAAGAAGACTACTGG 1749  
Qy 1240 aaagatggcatactacttatatggaggttctctcatcgtggtctcccccaatgcctcttc 1299  
Db 1750 AAAGAAGGTCCCTCGAGCTATATGGGTGTCCACGTTAAATTAACATATCCAAACATGTTCATG 1809  
Qy 1300 gtcgcacagggctcaagcgcgcgcgtctcttcaacgagcccaacgatcatagaacccaa 1359  
Db 1810 GTGCTTTGGACCAATATGGCCGCT-----TTACCAAGCTGCCGCCATCAATTAATGATACAG 1863  
Qy 1360 gtcgacttgatcgcatatcaaatgtgcaagttggagggcgagcagcgcacgtccgttgag 1419  
Db 1864 GTGGAATGATCACTGATACCAATTAATACAGGTTTGAACAAATGTTGAATCCATTGAA 1923  
Qy 1420 gcacgaatcagcagcagagaggtgctgatttatgattgacgaagatgaacagcacact 1479  
Db 1924 GCGACAAAAGAGCGGAAGACATGGACTCAAACTTGCGCCCAATATTTCGCGAANTGCC 1983

Qy 1480 ctgtcccttgacggattcgtggtgactggagggcaacatccctgggaagcaacacgt 1539  
Db 1984 TTATTCCCTTAAAGCGCAATCTCGGATTTTGGTCCGAATATCCGCGCAAGAAAACACG 2043  
Qy 1540 gctttaacccttcattagcgcg 1559  
Db 2044 GTTTACTTCTATCTCGGTGG 2063

RESULT 13  
AB026668  
LOCUS AB026668  
DEFINITION Acinetobacter sp. NCIMB9871 cyclohexanol metabolic gene cluster  
(chnB, chnE, chnF, chnG, orf3, chnA, orf5, chnD, chnH), complete cds.  
ACCESSION AB026668  
VERSION AB026668.1 GI:14587400  
KEYWORDS 6-hexanolactone hydrolase; 6-hydroxyhexanoate dehydrogenase;  
hypothetical 33.2 kDa protein; cyclohexanol dehydrogenase;  
hypothetical 36.6 kDa protein; xylS/araC-type transcriptional  
activator; 6-oxohexanoate dehydrogenase; cyclohexanone  
1,2-monooxygenase.

SOURCE Acinetobacter sp. NCIMB9871 (strain:NCIMB 9871) DNA.  
ORGANISM Acinetobacter sp. NCIMB9871  
Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
Acinetobacter.

REFERENCE 1 (bases 1 to 16160)  
AUTHORS Iwaki,H., Hasegawa,Y., Teraoka,M., Tokuyama,T. and Lau,P.C.  
TITLE Identification and Characterization of Cyclohexanol Metabolic Genes  
from Acinetobacter sp. NCIMB 9871  
JOURNAL Unpublished  
AUTHORS Hasegawa,Y. and Iwaki,H.  
TITLE Direct Submission  
JOURNAL Submitted (25-Apr-1999) Yoshie Hasegawa, Kansai University,  
Department of Biotechnology, Faculty of Engineering, Yamate-cho  
3-3-35, Suita, Osaka 564-8680, Japan  
(E-mail:yoshie@ipc.ku.kansai-u.ac.jp, Tel:81-6-6368-0909)  
LOCATION/Qualifiers  
1. .16160  
/organism="Acinetobacter sp. NCIMB9871"  
/strain="NCIMB 9871"  
/db\_xref="taxon:93373"  
538. .2169  
/gene="chnB"  
538. .2169  
/gene="chnB"  
/codon\_start=1  
/transl\_table=11  
/product="cyclohexanone 1,2-monooxygenase"  
/protein\_id="BAB61738.1"  
/db\_xref="GI:14587401"  
/translation="MSQKMFDAIVGGFGGLYAVKKLRDELELKVQAFDKATDVAG  
TWNRTFGALTOTETHLYCTSWDKELLSLEIKKRYQGPVDVVRKTLQQAERHDLRK  
SYQNTAVQSAHYHEADALVEVTEYDQKTYTARFLITLALGSLAPNLKIGNQPKG  
ELHITSRMPDVSFEKRVGVIGTGSTGVITAVAPLAKHLTVFORSQAYSVPIGND  
PLSESDYKTIKONYDKIWGVWNSALAFGLNESTVPAMSVSAERKAFERAKWTGGG  
PRMFTEFGDIATNMEANFIAQNFITKGLAEIVKDPAAIAQKLMQDLIAKRLQDSGY  
YNTFNRDNVLEDDKANPIVEITENGKLENGDEVELDMLICATGDAVDGNYVRMDI  
QKNGLAMDKYKRGSSYMGVTNNPMPVLGPNPPTNLPPSIESQVMSIDTI  
QYTVNNVESIEATKEAEQWQTQTCANTMTLPKQSWIFGANIPGKKNITVYFLG  
GLKYSALANCKNHAVEGFDIQLORSIDIKOPANA"

gene 2237. .3670  
/gene="chnE"  
2237. .3670  
/gene="chnE"  
/transl\_table=11  
/product="6-oxohexanoate dehydrogenase"  
/protein\_id="BAB61739.1"  
/db\_xref="GI:14587402"  
/translation="MNYNPILYINGEFLDHTNRDVKVEFNPVNHHCIGLMACASQAD  
LDVALESSQOAFLRWKKTSPTTRSEILRTFAKLAREKAAELIGNRIILDGKPLKEAIA







Db 5639 ACCGCGGTTCAATCGCTCATTTACAGGAAGCAGATGCTTGTGGGAAGTCACCACTGAA 5698  
QY 400 gctggccattgttgacgagcaagatcatcttcctacgaggttgcctccacgaag 459  
Db 5699 TATGGTGATAGTACAGCGCGCTTCCCTCATCACTGCTTTAGCGTTATTGTCTGCGCT 5758  
QY 460 cacactcccgactcccgcccttcgcgatttcaacggaaggtgattcattcgagtgc 519  
Db 5759 AACTGCGCAACATCAAGCAATTAATCAGTTTAAAGGTGAGCTGCATCATACACGCCG 5818  
QY 520 tggcaggaagcttcgacgagagggccagagagtcgcgtcagtcggtgcggggccaca 579  
Db 5819 TGCGCAGATGACGATGAGTTTGAAGGTAAACGCTGCGCGGTGATTTGTACGGGTCCACC 5878  
QY 580 agctccagagattgttcaggagttggccaaagaggtgcacagggttaacattgttatgca 639  
Db 5879 GGTGTTCAAGTTATTAGCGGTGTGGCAGCTCTGGCTAAACACCTCACTGCTTCCAGCGT 5938  
QY 640 agcgagagctattgtctgccatcgcgcaacgaacgattggtatggaacgaacagagcc 699  
Db 5939 TGTGCAATACAGCGTTCCCAATTTGCAATGATCCACTGCTCTGAAGAAGATGTTAAAG 5998  
QY 700 tgggaagggctactaccacagcgtgttgagcagagtcgaaagctctcgattgattccg 759  
Db 5999 ATCAAGAGCAATATGACAAATTTGGGATGCTGATGGAATTCAGCCCTTGCCTTTGCG 6058  
QY 760 gtccaggaacgctcggttgccatcttgaaagtcagcccgagcagcgagggagcctattc 819  
Db 6059 CTGAATGAAGCACAGTCCGACGATGAGCGTATCAGCTGAAGAAGCAAGCGCATGTTT 6118  
QY 820 gaagctgtggagcgtggggcctttaaattcttctgttccagtagccagagtcag 879  
Db 6119 GAAGAGCATGCGCAACAGTGGCGGTTCCTGTTTCATGTTTGAACTTTCGGTGATAT 6178  
QY 880 gtgcacaaagggccacccagctggctatgactcttggggccaaagagctcagctcgt 939  
Db 6179 GCCACATATGGAAGCCATATCGAAGCGCAAAATTCATTAAGGGTAAATTCGTGAA 6238  
QY 940 atgctcaatccggcaagagagatctcatggtctctctggagccgctgactggctcgt 999  
Db 6239 ATCGTCAAGATTCACG-----CATGTGCACAGAGCTTATGCCACAGGATTTGTAT 6289  
QY 1000 accagcgtctccactggagagcgactactacgaaatgctgggcaagcgagcgtcgaa 1059  
Db 6290 GCANAACGTCGCTGTGTGACAGTGGTTACTACACACCTTTAAACCGTGACAAATGTCCT 6349  
QY 1060 attgtgaatctagacaatcgcccatgtggtgttcaacagacaggtgtgctcttgagt 1119  
Db 6350 TTAGAAGATGTGAAGCGCAATCGGATTTGAAATTCACGAAACCGTGTGAACACTCGAA 6409  
QY 1120 gacgggagcaagagggaatgcacagcagctgtgctggcgaggggttcgacagttcact 1179  
Db 6410 AATGCGATTTCTGTTGAATTAAGATGCTGATATGTGCCACAGCTTTTGTATGCCGCTGAT 6469  
QY 1180 ggtcattgacacatatggcttgaacaaacagcagcagtgagtgacctgaagaggtgag 1239  
Db 6470 GGCACACTATGCGCATGGACATCAAGGTAAAGCGCTTGGCCATGAAGACTACTCG 6529  
QY 1240 aaagatgacatatctattatggagctctctctctctctggtctcccaatgctctctc 1299  
Db 6530 AAGAAGTTCGCTGACCTATATGGGTGACCGTAACTAACTATCCAAACATGTTCATG 6589  
QY 1300 gtgcacagcgtctgaagcccgacgcctcttccaaagcccaacgacatcatagaaaccaa 1359  
Db 6590 GTGTTGGACCGAATGGCCGT-----TTACCAACCTTGGCCGCGCATCAATGTAATCACAG 6643  
QY 1360 gtgcattgatcgccgatacaattgcaagattggagggcgagcgcagcgcctcgttgg 1419  
Db 6644 GTGGAATGGATCGATGATACCATTTCAATACCGGTTGAAACAATGTTGAATCATTTAA 6703  
QY 1420 ggcaggaatcagcacagagggatggtgattatgattgcacagatgaacagagcacact 1479

Db 6704 GCGAAAGAGCGGAACAACATGACTCAAACTTGCGCAATATTCGGGAATGACC 6763  
QY 1480 ctgttcccttgacgattcggtggactgagcgaacatccctgggaagcaacacagt 1539  
Db 6764 TTATTCCTTAAGCGCAATCTTGGAATTTTGGTGGATATTCGGGCAAGAAACACAG 6823  
QY 1540 gcttaaccccttcattagcgg 1559  
Db 6824 GTTACTTCTATCTCGGTGG 6843

RESULT 15  
AF257214 10549 bp DNA linear BCT 16-JUL-2001  
LOCUS  
DEFINITION  
Brevibacterium sp. HCU cyclohexanol and cyclohexanone oxidation  
gene cluster 1, complete sequence.  
ACCESSION  
AF257214  
VERSION  
AF257214.2 GI:14719371  
KEYWORDS  
SOURCE  
ORGANISM  
Brevibacterium sp. HCU.  
Brevibacterium sp. HCU  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Micrococciaceae; Brevibacteriaceae; Brevibacterium.  
REFERENCE  
AUTHORS  
1 (bases 4057 to 5718)  
Brzostowicz, P. C., Gibson, K. L., Thomas, S. M., Blasko, M. S. and  
Rouvriere, P. E.  
TITLE  
Simultaneous identification of two cyclohexanone oxidation genes  
from an environmental Brevibacterium isolate using mRNA  
differential display  
J. Bacteriol. 182 (15), 4241-4248 (2000)  
20353458  
10894733  
2 (bases 1 to 10549)  
Brzostowicz, P. C., Blasko, M. S. and Rouvriere, P. E.  
REFERENCE  
AUTHORS  
Identification of Two Gene Clusters Involved in Cyclohexanone in  
Brevibacterium sp. HCU  
3 (bases 4057 to 5718)  
Brzostowicz, P. C. and Rouvriere, P. E.  
TITLE  
Direct Submission  
Submitted (17-APR-2000) CR&D, DuPont, PO Box 80328 E328/B46,  
Wilmington, DE 19880-0328, USA  
4 (bases 1 to 10549)  
Brzostowicz, P. C. and Rouvriere, P. E.  
TITLE  
Direct Submission  
Submitted (29-DEC-2000) CR&D, DuPont, PO Box 80328 E328/B46,  
Wilmington, DE 19880-0328, USA  
REMARK  
Sequence update by submitter  
COMMENT  
On Jul 16, 2001 this sequence version replaced gi:9864057.  
FEATURES  
Location/Qualifiers  
source  
1..10549  
/organism="Brevibacterium sp. HCU"  
/strain="HCU"  
/db\_xref="taxon:133406"  
352..1728  
/gene="chxn1"  
352..1728  
/gene="chxn1"  
/codon\_start=1  
/transl\_table=11  
/product="putative sigma 54 type regulator"  
/protein\_id="AAK73159.1"  
/db\_xref="gi:14719372"

misc\_feature 1788..1916



```

/misc_feature
/notes="similar to Xanthomonas sp. transposon Tn5053
transposase encoded by GenBank Accession Number L20692"
complement(2510..2653)
/notes="similar to Mycobacterium tuberculosis probable
transposase encoded by GenBank Accession Number X65618"
3058..3877
/genes="cbnC1"
3058..3877
/genes="cbnC1"
/codon_start=1
/transl_except="(pos:3058..3070,aa:Met)
/transl_table=11
/product="caprolactone hydrolase"
/protein_id="AAK73160.1"
/db_xref="GI:14719373"
/translation="NPTTSSFSQSKQEVLTAPHTPTDPLGEIYAWEKDFREHPTM
SLQMRVPEDMQWRTKEPSNVRYEETEGSVFGIWLPEADDAKPLVLHGGFAL
GSNSRKLKAGHAKQSGRQAFVADFLPAEPHPPAQIEDALTVISAMNSRGIPFENI
TLVGDGASGASIAIGTVLSLLKDGRLPRQVVTMPWMDMSNGSETIESNDAYDFLTR
DGLQGNIDRYLAVERILTVGLVNPFLYADFHGFPRLYICVSDTESSTRASV"
4057..5718
/genes="cbnB1"
4057..5718
/genes="cbnB1"
/EC_number="1.14.13.22"
/notes="flavin-dependent monooxygenase; contains FAD;
requires NADPH"
/codon_start=1
/transl_table=11
/product="cyclohexanone monooxygenase 1"
/protein_id="AAG01289.1"
/db_xref="GI:9864058"
/translation="MPITQQLDHDIAIVIGAFSGLSAILHLHREILGRQIVEATDGIG
KTIQNRVPEVDTSEPHYSFSEKVRDEWWTQRYPDGEEVCAYLNFADRLDLR
GDQLMSRVNTARMNETKYWDVIFEDGSKRARFLISALGASQAIFFADIGDEFN
GAKYHTAAQVDCVTCGKVGIVGASGIOIIPELAKAGELFVORTPNYVVEFN
NDKVAEMQVYRDNYDEIFERASKPGVDMEYDTSADAVESEERKVSFEKWEKG
GFHFNCECTDLGTSPSEASELASEFIRSKREVVDKPDADILLCPKSYFNGKRVPTG
HGYETFNNTVLLDQARTPIRISKGIHVHGDTEYELDAIVFATGDMATGTLTNI
DIVRGVGLTRKMAQDGLRTNIGLTVNGFNFELMSLGPQTPYSNLVPEIOLGAWMO
RFLKFLQERGIETVSEAREEINNAETIRGAESTVMSLEGPKAGAWTIGNIFGKSR
EYQVYMGQGVQVDCRBAEESDIATFLNADSIDGEKVRASAGK"
5832..7096
/genes="cbnD1"
5832..7096
/genes="cbnD1"
/notes="6-hydroxyhexanoate dehydrogenase"
/codon_start=1
/transl_table=11
/product="Fe-dependent alcohol dehydrogenase"
/protein_id="AAK73161.1"
/db_xref="GI:14719374"
/translation="MESHNENTLGLLRQGVTVVFGPGRRELPSIARRYGSTVLIC
TDERMLAEPWCIDLOTALEKAGMRVYVGNVRPDLPRADIQTATRLKLAHDKIDVIFGL
GGSCDMFAKVMGILLSPGVDRIFGENVYSGPGLPVITVPTTGTGAETACLSVYH
DEKGVKGVASAYMQAVTVIDPEFTLTAPEGLTAATADALSHLVSYTAYKNPS
SDIRDLHYGKNLLTDYWAERGLKLLSDGIPALAKDLTDLNARTNMVLAAPFCGKGI
NTWTAGCALQSPLSALTCTSHGPGYCALLPYMYBNLNPASTPFAELGELVGDRC
STYLESQHNKXKVEWLYSTIGAPTDLGALGTEADVAGVAKAAASFLIANPERPL
PALIMEILLNVRGRDSWHAEE"
7084..8034
/codon_start=1
/transl_table=11
/product="hydroxacyl-CoA dehydrogenase"
/protein_id="AAK73162.1"
/db_xref="GI:14719375"
/translation="MGRVSHGVIFGAGSIGTAFALLFADAGFAVRIFDPDPSALERSR
HVIDQRTILQRLTLLASNPSEVRELIEIVSSARTASGAILVQEAGPEDVQTKHIF
EDLTVTSETILASASAIIPSSRFVDVHSAFSLIGHPCNPPIYLRVVELVGNPSTE
EQTLRAGOLYEQAGLSAVRVNRVDGVFNRIQGVAREVALYVGBIIVDMOLDIL
VODGLRNSVAGPFATVDLNVRRGGITAHARMGSAYHRMAGALDTSKEDWDTLVAKV
NCSRRKAVLEQWDQVADRDTQLMKQLNARTSNGTTRD"
8018..8794

```

```

/codon_start=1
/transl_except="(pos:8018..8020,aa:Met)
/transl_table=11
/product="short-chain dehydrogenase"
/protein_id="AAK73163.1"
/db_xref="GI:14719376"
/translation="MLPVTDSLGGDVLVTGGAGIGKATTTALAEGRGRVVLTDVDE
DAGSVADEVRENTGETRFPDLVTPAAVTECAOKLDDEGMPVYGLMANAGTAPSS
SANDVDELRLSTVDINLNGVFCRFGKRMILARGSVVTTSSIAGRPTVSPERHA
AYGATKAANVHLVGLLVETWAKTGVNVAVAGYRTPTILEALKARESETISEWTERI
PAGRLNDPSRIADGVVFLMSNARKIITGVTLHDGGYAAR"
9132..9902
/genes="cbnA"
9132..9902
/genes="cbnA"
/notes="similar to short-chain dehydrogenase family
proteins"
/codon_start=1
/transl_table=11
/product="cyclohexanol dehydrogenase"
/protein_id="AAK73164.1"
/db_xref="GI:14719377"
/translation="MNRLLGKRVAVITGTAAGMGRIOSELYASEGAQVAVVDVNBQGR
ATADAJRASSGVANYKLDVSDSEVELVSDIAKRFAGINLVNAGVTGADKPTHE
IDERDLDLVSLVDVKGVEFMKHCIPYEKOAGGSAIVNEASTYGLVGSQELTPVHAAK
GAVVALTKODAVTYGPSNIRVNAVAPGTLITPLVKELSGSRGPDGLDGYTKLWAKHPL
GRVGTPEEVAANTPLASEASEASFTIGAVLVPVDDGGYTAQ"
complement(9934..10341)
/notes="similar to para-aminobenzoic acid synthase from
Streptomyces encoded by GenBank Accession Number U60417"
BASE COUNT 2528 a 2895 c 2842 g 2284 t
ORIGIN

Query Match 8.6%; Score 140.6; DB 1; Length 10549;
Best Local Similarity 47.2%; Pred. No. 3.4e-25;
Matches 567; Conservative 0; Mismatches 619; Indels 15; Gaps 4;

QY 45 CQACQCCATCATCTGTAGCCGCGCTCAGCGGCTCTCTCTGTGTACAAATGCGAA 104
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4083 CGACGCTATGCTCATCGCGCGCGCTTCTCCGGAGTAGCCATTCTGCACCACTGCGTGA 4142
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 GTCAGACTCAACCAATCTCAGGAGAGCCGCCCTATTTCGCGGCTCTGGCACTG 164
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4143 AATCGCCTAGACACTCAATTCGTGAGCACCCAGCGCATTCGAGGAATTTGGTGGAT 4202
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 GAACCGTACCTGGCGCTGCTGTGATCGAGAGCGCCCTCTTACCACTGAACATTC 224
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4203 CAACCGCTACCGCGGGTGGCGACCGAGTTCACACTACTCTTTTCAGCTTCAG 4262
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 C---GAAGTATGAAAGACTGGACCTGGTCTGCGCTATCTCCTGACCAAGAGTGTG 281
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4263 CAAGGAAGTTGTCGACGAGTGGACATGGACTCAACGCTACCCACAGCGGTGAAGAAGTTTG 4322
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 GTCATATGTCACCACTGACAAGTCCGCGGCTTGAGAAAGACGCTCCTACTTCGAGC 341
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4323 CGCCTATCTCAATTCATTGCTGATCGACTTGATCTTCCGGAAGGACATTCAGCTCAACTC 4382
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 342 TGAAGTGTGATCGCGGATGCCAGAGATCTGGGCACTGGACTGTCAAGACGCTCGGC 401
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4383 ACGAGTGAATACTCCCGCTGGGAATGAGCGGNAAGTACTGGGACGCTATTTTCGAGA 4442
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 402 TGGCCATGTCGAGCGGCAAGTATCTCCTGCTACGCGGCTGTCCACAGGAAGCA 461
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4443 CGGGTCTCGAAGACGCGCTCGCTCTCTCATACGCGCATGGGTGCACATTCACCGAGCGAT 4502
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 462 CACTCCCGCACTCCCGCGCTCCCGCTATTTCACACGGAAGTGTGATTCATTCAGTGCCTG 521
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4503 TTTCCCGGCGCATTCGACGGAATTCACGAATTCACACGCGCGCAAAATATCACATCGCGCTT 4562
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 522 GC---ACGAAGACTTCACGAGGCGGCGAGAGTCGCGCTCATCGGTGCGGCGGCGCAC 578
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4563 GCCAGCTGATGCGCTAGATTTCACGGGCAAGAAGTTGAGAGTCATTGGGTTGGGCGCTC 4622
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```





Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1238 ggaagatgcatatctatt 1260  
 Db 260 GGAAGATGCATATCTATTAT 238

## RESULT 2

LOCUS B045552/c 642 bp mRNA linear EST 05-DEC-2001  
 DEFINITION B045552 NTBB Mochil normalized Xenopus neurula library Xenopus  
 laevis cDNA clone XL004115 3', mRNA sequence.

ACCESSION B045552  
 VERSION B045552.1 GI:17369603

KEYWORDS EST.  
 SOURCE African clawed frog.

ORGANISM Xenopus laevis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

AUTHORS Kitayama, A., Terasaka, C., Mochil, M., Ueno, N., Shin-i, T. and Kohara

Y.

TITLE Expressed genes in X. laevis embryo

JOURNAL Unpublished (2001)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshinigenes.nig.ac.jp.

FEATURES Location/Qualifiers

1..642

/organism="Xenopus laevis"

/db\_xref="taxon:8355"

/clone\_lib="NIBB Mochil normalized Xenopus neurula

library"

/tissue\_type="whole embryo"

/dev\_stage="stage 15"

153 a 152 c 122 g 214 t 1 others

BASE COUNT 153 a 152 c 122 g 214 t

ORIGIN

Query Match 1.4% Score 23; DB 10; Length 642;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1238 ggaagatgcatatctatt 1260

Db 368 GGAAGATGCATATCTATTAT 346

RESULT 3

LOCUS AM478042 232 bp mRNA linear EST 25-APR-2001

DEFINITION 18283 MARC lBOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AM478042

VERSION AM478042.1 GI:7048148

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 232)

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett

G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G.,

Pertea, G., Holt, L., Karamycheva, S., Liang, F., Quackenbush, J. and

Keeler, J.W.

libraries and construction of a gene index for cattle

Genome Res. 11 (4): 626-630 (2001)

21180013

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 20

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCACAGTCAGCAGC

Plate: 12 row: L column: 3

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1..232

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC lBOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;

Library made from pooled tissue from lymph node, ovary,

fat, hypothalamus, and pituitary."

84 a 38 c 31 g 79 t

BASE COUNT 84 a 38 c 31 g 79 t

ORIGIN

Query Match 1.2% Score 20; DB 9; Length 232;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1089 ggctttacaaagacaggtg 1108

Db 108 GGGCTTTACAAAGACAGGTG 127

RESULT 4

LOCUS AQ708508 451 bp DNA linear GSS 07-JUL-1999

DEFINITION HS\_5369\_A2\_A04\_SP6E RPCI-11 Human Male BAC library Homo sapiens

genomic clone Plate-945 Col-8 Row-A, DNA sequence.

ACCESSION AQ708508

VERSION AQ708508.1 GI:5417934

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 451)

Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., and

Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and

HOOD, L.

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

99380589

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

## JOURNAL

## MEDLINE

## COMMENT

plate: 945 row: A column: 8  
Seq primer: SP6  
Class: BAC ends  
High quality sequence stop: 451.  
Location/Qualifiers

## FEATURES

source

1. .451  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate-945 Col-8 Row-A"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/sex="male"  
/note="Vector: pRAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRV. Methylation. Size selected DNA was cloned into the  
pRAC3.6 vector at EcoRI sites"

BASE COUNT

145 a 65 c 90 g 151 t

ORIGIN

Query Match

1.2%; Score 20; DB 12; Length 451;

Best Local Similarity

100.0%; Pred. No. 27;

Matches

20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 attatgattgcgaatgaa 1469

|||||

Db 156 ATTATGATGCGCAAGATGAA 175

RESULT 5

AV644438/c

LOCUS

AV644438 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii

DEFINITION

CDNA clone HCL088e07\_r 5', mRNA sequence.

ACCESSION

AV644438

VERSION

AV644438.1

KEYWORDS

EST.

SOURCE

Chlamydomonas reinhardtii.

ORGANISM

Chlamydomonadaceae; Chlamydomonas.

REFERENCE

1 (bases 1 to 544)

AUTHORS

Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohyama K.,

Matsumura Y. and Tabata S.

TITLE

Generation of expressed sequence tags from low-CO2 and high-CO2

JOURNAL

adapted cells of Chlamydomonas reinhardtii

MEDLINE

DNARES. 7 (5), 305-307 (2000)

COMMENT

20539644

Contact: Erika Asamizu

The First laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizukazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.

Location/Qualifiers

1. .544

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db\_xref="taxon:3055"

/clone\_lib="HCL088e07\_r"

/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:

XhoI; The cDNA library was constructed from cells cultured

in a medium with bubbling air containing 5% carbon

dioxide"

BASE COUNT

103 a 177 c 161 g 102 t 1 others

ORIGIN

Query Match

1.2%; Score 20; DB 9; Length 544;

Best Local Similarity

100.0%; Pred. No. 28;

Matches

20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 ttgagccgcgcctcagcggc 78

Db 499 TTGGAGCGCGCCTCAGCGC 480  
|||||

RESULT 6

BG685227

LOCUS

BG685227

DEFINITION

602637125F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4764582 5',

mRNA sequence.

ACCESSION

BG685227

VERSION

BG685227.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 788)

AUTHORS

NTH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: graphs@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LNCMI622 row: b column: 07

High quality sequence stop: 651.

FEATURES

Location/Qualifiers

1. .788

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4764582"

/clone\_lib="NIH\_MGC\_48"

/tissue\_type="primary B-cells from tonsils (cell line)"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;

Site\_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected &gt;500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

BASE COUNT

230 a 163 c 198 g 197 t

ORIGIN

Query Match

1.2%; Score 20; DB 10; Length 788;

Best Local Similarity

100.0%; Pred. No. 31;

Matches

20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 384 gactgtcaagcgcgcgcgtg 403

|||||

Db 733 GACTGTCAAGACGTCGCGCTG 752

RESULT 7

AV291651/c

LOCUS

AV291651

DEFINITION

AV291651

ACCESSION

AV291651

VERSION

AV291651.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 240)

AUTHORS

RIKEN full-length enriched, 6 days neonate head Mus

TITLE

musculus cDNA clone 5430408H04 3', mRNA sequence.

JOURNAL

Unpublished (1999)

COMMENT

Contact: Erika Asamizu

Email: asamizukazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.

Location/Qualifiers

1. .544

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db\_xref="taxon:3055"

/clone\_lib="HCL088e07\_r"

/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:

XhoI; The cDNA library was constructed from cells cultured

in a medium with bubbling air containing 5% carbon

dioxide"

BASE COUNT

103 a 177 c 161 g 102 t 1 others

ORIGIN

Query Match

1.2%; Score 20; DB 9; Length 544;

Best Local Similarity

100.0%; Pred. No. 28;

Matches

20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 ttgagccgcgcctcagcggc 78

REFERENCE  
AUTHORS

1 (bases 1 to 240)

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,Y., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateo,M., Tominaga,N., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Konno,H., et al. 1999)  
Unpublished (1999)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.  
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7): 3455-3460 (1998)  
Itoh,M., Kitsuami,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5): 463-470 (1999)  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES  
source

Location/Qualifiers  
1..240  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="5430408H04"  
head  
/sex="mixed"  
/tissue.type="head"  
/dev\_stage="6 days neonate"  
/lab\_host="DH10B"  
/note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15'.  
GAGAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTNN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adaptor of sequence 15'. GAGAGAGATTCGAGTTAATTAATTCCTCCCCCCC 3'. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI, 3' end: BamHI."

BASE COUNT  
ORIGIN

79 a 42 c 20 g 99 t

## Query Match

Best Local Similarity 1.2%; Score 19; DB 9; Length 240; Mismatches 0; Conservative 100.0%; Pred. NO. 79;

## Matches

19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1058

aaattgtgaattagaaaca 1076

Db 20

AAATTGTGAATTAGAAC A 2

## RESULT

8

AQ485204

LOCUS

RPCI-11-233G14-TV RPCI-11 Homo sapiens genomic clone RPCI-11-233G14

DEFINITION

, DNA sequence.

ACCESSION

AQ485204

VERSION

AQ485204.1 GI:4667252

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 321)

AUTHORS

Zhao,S., Adams,M.D., Nierman,M., Malek,J., de Jong,P. and Venter,J.C.

TITLE

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL

Unpublished (1997)

COMMENT

Other GSSs: RPCI-11-233G14.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pdejong@med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

Research Genet cs (info@resgen.com). BAC end search page:

[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).

Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers

1..321

/organism="Homo sapiens"

/db\_xref="GDB:7589245"

/db\_xref="taxon:9606"

/clone="RPCI-11-233G14"

/sex="Male"

/cell\_type="Lymphocytes"

/notes="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;

RPCI11 Human Male BAC Library"

BASE COUNT

66 a 86 c 79 g 90 t

ORIGIN

Query Match

Best Local Similarity 1.2%; Score 19; DB 12; Length 321;

Matches

19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 417

ggcaagatctctcttc 435

Db 14

GGCAAGATCTCTCTCTC 32

## RESULT

9

BB451655/C

LOCUS

BB451655 RIKEN full-length enriched, 12 days embryo spinal ganglion

DEFINITION

Mus musculus neuritin mRNA, mRNA sequence.

ACCESSION

BB451655

VERSION

BB451655.1 GI:9312690

KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS

1. Endo, T., Fukuda, S., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Hirazane, T., Hori, F., Ishii, F., Ishikawa, J., Ishikawa, T., Itoh, M., Iizawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Konno, H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suihicho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermotabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Kikunai, T., Akiyama, J., Shibata, K., Iizawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES  
source

1..327  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="0130012C11"  
/clone\_lib="RIKEN full-length enriched, 12 days embryo  
spinal ganglion"  
/tissue\_type="spinal ganglion"  
/dev\_stage="12 days embryo"  
/lab\_host="DH10B"  
/note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCACTCGAGTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGATTAAATTAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I." 108 a 75 c 42 g 102 t

Query Match 1.2%; Score 19; DB 9; Length 327;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1216 ggaatggacctgaaggagg 1234  
|||||  
Db 78 GCGTGGACCTGAGAGAGG 50

RESULT 10  
LOCUS AW747522/c  
DEFINITION WSI\_68\_H02\_g1\_A002 Water-stressed 1 (WS1) sorghum bicolor cDNA, mRNA sequence.  
ACCESSION AW747522  
VERSION AW747522.1 GI:7661260  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 340)  
AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L. H.  
TITLE An EST database from Sorghum: water-stressed plants  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonnier-Pratt MM  
Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 542 1805  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: T7  
High quality sequence stop: 329  
POLYA=yes.  
Location/Qualifiers  
1..340  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Water-stressed 1 (WS1)"  
/note="Organ: Mix of 3-week old plants on days 7 & 8 after water was withheld; Vector: Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was made from polyA RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision." 78 a 76 c 88 g 98 t

BASE COUNT  
ORIGIN

Query Match 1.2%; Score 19; DB 9; Length 340;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 578 caagcatccagatgttcca 596  
|||||  
Db 164 CAAGCATCCAGATTGTCCA 146

RESULT 11  
LOCUS A1526742  
DEFINITION uJ2d02.y1 Sugano mouse liver m11a Mus musculus cDNA clone IMAGE:1942595 5' similar to gb:M73490 Mus musculus apolipoprotein E mRNA, 3' end (MOUSE); mRNA sequence.  
ACCESSION A1526742  
VERSION A1526742.1 GI:4440877  
KEYWORDS EST.  
SOURCE house mouse.





Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: cc@resgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 427.  
 Location/Qualifiers  
 1. .428

## FEATURES

source

/organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1074-4504"  
 /clone\_lib="Gm-c1074"  
 /tissue\_type="seedlings induced for HR (hypersensitive  
 response)"  
 /dev\_stage="9-11 day old"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI. The cDNA library was constructed from mRNA isolated  
 from 9-11 day old seedlings that were induced for HR  
 (hypersensitive response) by vacuum infiltrating plant  
 tissue with *Pseudomonas syringae* pv. *glyciniae* carrying the  
 avrB gene (Genetics 141:1157-1604). Plant tissue (expanded  
 unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36  
 and 53 hrs after inoculation and their mRNA pooled equally  
 for cDNA construction. The library was prepared using the  
 Stratagene pBluescript II SK(+) library construction kit.  
 Complementary DNA was synthesized from mRNA using a primer  
 consisting of a poly(dT) sequence with an XhoI restriction  
 site. EcoRI adapters were ligated to the blunt-ended cDNA  
 fragments followed by XhoI digestion. The cDNA insert is  
 protected from XhoI digestion via methylation during first  
 strand synthesis. The cDNA fragments were directionally  
 cloned into the EcoRI-XhoI restriction site of the  
 pBluescript vector. The ligated cDNA fragments were  
 transformed into *E. coli* Electromax DH10B host cells. Plant  
 care, inoculations, and library construction were  
 performed by Steve Clough (Lila Vodkin lab, University of  
 Illinois)."

BASE COUNT 84 a 85 c 71 g 109 t 79 others  
 ORIGIN

Query Match 1.2% Score 19; DB 10; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 ctaccaactgaacattccc 225  
 |||||  
 Db 113 CTACCACTGAACATTCCC 95

RESULT 14  
 TA151G030/c  
 LOCUS  
 DEFINITION T. brucei sheared genomic DNA clone 151g03, reverse sequence,  
 genomic survey sequence.  
 ACCESSION AL473065  
 VERSION AL473065.1 GI:11838338  
 KEYWORDS GSS.  
 SOURCE Trypanosoma brucei.  
 ORGANISM  
 Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma.  
 1 (bases 1 to 435)  
 REFERENCE  
 AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,  
 Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,  
 Melville, S.E., Rajandream, M.A. and Barrell, B.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing  
 project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and  
 nh@sanger.ac.uk  
 Constructed at the Institute for Genomic Research (TIGR),  
 Rockville, MD. Genomic DNA isolated from a cloned population of  
 Trypanosoma brucei (TR9U97/4 GUTat 10.1) was mechanically sheared  
 to give a tight size distribution (4 kb). The v + i method used for the library construction is  
 described in detail in Smith, H. and Venter, J.C. (Making small  
 insert libraries for whole genome shotgun sequencing projects. In  
 genome sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Barrell, Oxford University Press, 1999).  
 Email: nelsayed@tigr.org  
 Details of T. brucei sequencing at the Sanger Centre are available  
 at <http://www.sanger.ac.uk/projects/T-brucei/>.

## FEATURES

source

Location/Qualifiers  
 1. .435  
 /organism="Trypanosoma brucei"  
 /strain="TR9U97"  
 /db\_xref="taxon:5691"  
 /clone="151g03"

BASE COUNT 89 a 125 c 87 g 134 t  
 ORIGIN

Query Match 1.2% Score 19; DB 12; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 atcagcacagagcgtgg 1446  
 |||||  
 Db 99 ATCAGCACAGAGCGATGG 81

## RESULT 15

AT149258/c

LOCUS

DEFINITION

qk53c03.x1 NCI\_CGAP\_Panl Homo sapiens cDNA clone IMAGE:2005060 3',

similar to gb:V00567 BETA-2-MICROGLOBULIN PRECURSOR (HUMAN);, mRNA

sequence.

ACCESSION AT149258

VERSION AT149258.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 440)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

unknown library type

Insert Length: 675 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 267.

Location/Qualifiers

1. .440

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2005060"

/clone\_lib="NCI\_CGAP\_Panl"

/tissue\_type="adenocarcinoma"

/lab\_host="DH10B"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.72 kb. Life Technologies catalog #:

11348-013

BASE COUNT 150 a 87 c 64 g 139 t

ORIGIN

Query Match 1.2%; Score 19; DB 9; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 442 ggattactccacaggaac 460  
 |||||  
 DB 361 GGGTTGCTCCACAGGAGC 343

Search completed: April 27, 2002, 02:58:36  
 Job time: 13428 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 02:59:08 ; Search time 95.74 Seconds  
(without alignments)  
3142.907 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638

Sequence: 1 atgtcggccaccagcaactc.....tgcttcgtctccctgctaa 1638

Scoring table:

OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 139433 seqs, 91850399 residues

Word size : 0

Total number of hits satisfying chosen parameters: 278866

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents\_NA\_New.\*

- 1: /cgn2.6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2.6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2.6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*
- 4: /cgn2.6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2.6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2.6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2.6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	19	1.2	21100	6	US-10-105-299-15122
2	18	1.1	216	5	US-09-540-210B-3305
3	18	1.1	261	5	US-09-975-254-28990
4	18	1.1	273	5	US-09-540-210B-32635
5	17	1.0	253	5	US-09-975-254-8723
6	17	1.0	263	5	US-09-975-254-13242
7	17	1.0	313	5	US-09-975-254-17647
8	17	1.0	531	1	PCT-US02-05706-38
9	17	1.0	681	1	PCT-US02-05706-26
c 10	17	1.0	1398	5	US-09-540-209B-4551
11	17	1.0	4874	6	US-10-116-802-157
12	17	1.0	7182	6	US-10-116-802-207
13	17	1.0	9702	6	US-10-105-299-13829
14	17	1.0	27976	6	US-10-105-299-13828
15	16	1.0	204	5	US-09-540-210B-21427
c 16	16	1.0	279	6	US-10-112-699-559
c 17	16	1.0	280	6	US-10-112-699-632
c 18	16	1.0	286	5	US-09-540-210B-27252
c 19	16	1.0	288	6	US-10-105-299-7116
c 20	16	1.0	288	6	US-10-105-299-7117
c 21	16	1.0	288	6	US-10-105-299-7312
c 22	16	1.0	288	6	US-10-105-299-7313
c 23	16	1.0	288	6	US-10-105-299-8633
c 24	16	1.0	288	6	US-10-105-299-8634
c 25	16	1.0	288	6	US-10-105-299-10321
c 26	16	1.0	288	6	US-10-105-299-10322

c 27	16	1.0	288	6	US-10-105-299-14375	Sequence 14375, A
c 28	16	1.0	288	6	US-10-105-299-14380	Sequence 14380, A
c 29	16	1.0	380	6	US-10-103-295-380	Sequence 380, App
c 30	16	1.0	397	6	US-10-103-295-374	Sequence 374, App
c 31	16	1.0	397	6	US-10-105-299-8488	Sequence 8488, App
c 32	16	1.0	405	6	US-10-112-699-724	Sequence 724, App
c 33	16	1.0	405	6	US-10-103-295-367	Sequence 367, App
c 34	16	1.0	414	6	US-10-105-299-8481	Sequence 8481, App
c 35	16	1.0	434	6	US-10-103-295-376	Sequence 376, App
c 36	16	1.0	434	6	US-10-105-299-8490	Sequence 8490, App
c 37	16	1.0	462	6	US-10-103-295-363	Sequence 363, App
c 38	16	1.0	462	6	US-10-105-299-8477	Sequence 8477, App
c 39	16	1.0	462	6	US-10-105-299-14700	Sequence 14700, A
c 40	16	1.0	509	6	US-10-103-295-378	Sequence 378, App
c 41	16	1.0	509	6	US-10-105-299-8204	Sequence 8204, App
c 42	16	1.0	509	6	US-10-105-299-8492	Sequence 8492, App
c 43	16	1.0	509	6	US-10-105-299-9177	Sequence 9177, App
c 44	16	1.0	509	6	US-10-105-299-10603	Sequence 10603, A
c 45	16	1.0	509	6	US-10-105-299-14713	Sequence 14713, A

#### ALIGNMENTS

RESULT 1  
US-10-105-299-15122/c  
; Sequence 15122, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15122  
; LENGTH: 21100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-15122

Query Match 1.2%: Score 19; DB 6; Length 21100;  
Best Local Similarity 100.0%; Pred. No. 0.71; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0;

Qy 689 aacagacagcctggaagc 707  
Db 5304 AACAGACAGCCTGGAAGC 5286  
|||||

RESULT 2  
US-09-540-210B-3305  
; Sequence 3305, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Deleageane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullaly, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526

```

; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/803,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/803,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997

```

```

; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 3305
; LENGTH: 216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00047457
US-09-540-210B-3305

Query Match          1.1%  Score 18;  DB 5;  Length 216;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy 1586 aagagaaggtggcccaatt 1603
    |||||||
Db 23 aagagaaggtggcccaatt 40

RESULT 3
US-09-975-254-28990
; Sequence 28990, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(15309)B
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/975,254
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 28990
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700899231H1
US-09-975-254-28990

Query Match          1.1%  Score 18;  DB 5;  Length 261;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy 1217 gaatgaacctgaaggagg 1234
    |||||||
Db 90 gaatgaacctgaaggagg 107

RESULT 4
US-09-540-210B-32635
; Sequence 32635, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 C1P

```

```

; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199

```

```

; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 32635
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: hu01324446
US-09-540-210B-32635

Query Match          1.1%; Score 18; DB 5; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1586 aagagaagtggtgccaatt 1603
      |||||
Db 193 aagagaagtggtgccaatt 210

RESULT 5
US-09-975-254-8723
; Sequence 8723, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Hyrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 8723
; LENGTH: 253
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700754933H1
US-09-975-254-8723

Query Match          1.0%; Score 17; DB 5; Length 253;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1291 gccttcttcgtcgccac 1307
      |||||
Db 62 gccttcttcgtcgccac 78

RESULT 6
US-09-975-254-13242
; Sequence 13242, Application US/09975254

```

; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Heck, Gregory R.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15309)B  
; CURRENT APPLICATION NUMBER: US/09/975,254  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US/09/263,191  
; PRIOR FILING DATE: 1999-03-05  
; NUMBER OF SEQ ID NOS: 31255  
; SEQ ID NO 13242  
; LENGTH: 263  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 700907418H1  
US-09-975-254-13242

Query Match 1.0%; Score 17; DB 5; Length 263;  
Best Local Similarity 100.0%; Pred. No. 9.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1268 tcttctctcattggttc 1284  
Db 92 tcttctctcattggttc 108

RESULT 7  
US-09-975-254-17647  
; Sequence 17647, Application US/09975254  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Heck, Gregory R.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15309)B  
; CURRENT APPLICATION NUMBER: US/09/975,254  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US/09/263,191  
; PRIOR FILING DATE: 1999-03-05  
; NUMBER OF SEQ ID NOS: 31255  
; SEQ ID NO 17647  
; LENGTH: 313  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 700954584H1  
US-09-975-254-17647

Query Match 1.0%; Score 17; DB 5; Length 313;  
Best Local Similarity 100.0%; Pred. No. 9.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 7 gccaccagcaactccag 23  
Db 253 gccaccagcaactccag 269

RESULT 8  
PCT-US02-05706-38  
; Sequence 38, Application PC/TUS0205706  
; GENERAL INFORMATION:  
; APPLICANT: Rohm and Haas Company  
; APPLICANT: Palli, Subba R.  
; APPLICANT: Kapitskaya, Marianna Z.  
; TITLE OF INVENTION: Chimeric retinoid x receptors and their use in a novel ecdysone  
; TITLE OF INVENTION: Based inducible gene expression system  
; FILE REFERENCE: A01238  
; CURRENT APPLICATION NUMBER: PCT/US02/05706  
; CURRENT FILING DATE: 2002-02-20

; PRIOR APPLICATION NUMBER: US 60/294,819  
; PRIOR FILING DATE: 2001-05-31  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 38  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Apis mellifera  
PCT-US02-05706-38

Query Match 1.0%; Score 17; DB 1; Length 531;  
Best Local Similarity 100.0%; Pred. No. 9.7; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1147 atcgtgctggcgacggg 1163  
Db 115 atcgtgctggcgacggg 131

RESULT 9  
PCT-US02-05706-26  
; Sequence 26, Application PC/TUS0205706  
; GENERAL INFORMATION:  
; APPLICANT: Rohm and Haas Company  
; APPLICANT: Palli, Subba R.  
; APPLICANT: Kapitskaya, Marianna Z.  
; TITLE OF INVENTION: Chimeric retinoid x receptors and their use in a novel ecdyson  
; TITLE OF INVENTION: Based inducible gene expression system  
; FILE REFERENCE: A01238  
; CURRENT APPLICATION NUMBER: PCT/US02/05706  
; CURRENT FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: US 60/294,819  
; PRIOR FILING DATE: 2001-05-31  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Apis mellifera  
PCT-US02-05706-26

Query Match 1.0%; Score 17; DB 1; Length 681;  
Best Local Similarity 100.0%; Pred. No. 9.7; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1147 atcgtgctggcgacggg 1163  
Db 265 atcgtgctggcgacggg 281

RESULT 10  
US-09-540-209B-4551/c  
; Sequence 4551, Application US/09540209B  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,1001-001  
; CURRENT APPLICATION NUMBER: US/09/540,209B  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 4551  
; LENGTH: 1398  
; TYPE: DNA  
; ORGANISM: B. fragilis  
US-09-540-209B-4551

Query Match 1.0%; Score 17; DB 5; Length 1398;  
Best Local Similarity 100.0%; Pred. No. 9.8; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;

QY 1269 cttctctcatgcttcc 1285  
|||||  
Db 1102 CTTCTCTCATGCTTCC 1086

RESULT 11  
US-10-116-802-157  
; Sequence 157, Application US/10116802  
; GENERAL INFORMATION:  
; APPLICANT: Amy Lasek  
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER  
; FILE REFERENCE: PA-0045 US  
; CURRENT APPLICATION NUMBER: US/10/116,802  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 60/281,593  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 519  
; SOFTWARE: PERL Program  
; SEQ ID NO 157  
; LENGTH: 4874  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: incyte ID No: 1330186CB1  
; NAME/KEY: unsure  
; LOCATION: 4822  
; OTHER INFORMATION: a, t, c, g, or other  
US-10-116-802-157

Query Match 1.0%; Score 17; DB 6; Length 4874;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 gcgcgcctctctgctgtg 90  
|||||  
Db 1578 gcgcgcctctctgctgtg 1594

RESULT 12  
US-10-116-802-207  
; Sequence 207, Application US/10116802  
; GENERAL INFORMATION:  
; APPLICANT: Amy Lasek  
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER  
; FILE REFERENCE: PA-0045 US  
; CURRENT APPLICATION NUMBER: US/10/116,802  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 60/281,593  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 519  
; SOFTWARE: PERL Program  
; SEQ ID NO 207  
; LENGTH: 7182  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: incyte ID No: 1099945.13  
US-10-116-802-207

Query Match 1.0%; Score 17; DB 6; Length 7182;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 atcatcgttgagccgg 68  
|||||  
Db 5287 atcatcgttgagccgg 5303

RESULT 13  
US-10-105-299-13829  
; Sequence 13829, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13829  
; LENGTH: 9702  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-13829

Query Match 1.0%; Score 17; DB 6; Length 9702;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 gggggccacaagcatcca 587  
|||||  
Db 7461 gggggccacaagcatcca 7477

RESULT 14  
US-10-105-299-13828  
; Sequence 13828, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13828  
; LENGTH: 27976  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-13828

Query Match 1.0%; Score 17; DB 6; Length 27976;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 gggggccacaagcatcca 587  
|||||  
Db 7461 gggggccacaagcatcca 7477

RESULT 15  
US-09-540-210B-21427  
; Sequence 21427, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; PRIOR FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244

; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997  
; PRIOR APPLICATION NUMBER: 60/014,010  
; PRIOR FILING DATE: March 25, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/903,555  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/862,178  
; PRIOR FILING DATE: May 22, 1997  
; PRIOR APPLICATION NUMBER: 60/018,217  
; PRIOR FILING DATE: May 23, 1996  
; PRIOR APPLICATION NUMBER: 08/881,589  
; PRIOR FILING DATE: June 24, 1997  
; PRIOR APPLICATION NUMBER: 60/021,275  
; PRIOR FILING DATE: June 25, 1996  
; PRIOR APPLICATION NUMBER: 08/903,802  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/905,881  
; PRIOR FILING DATE: August 1, 1997  
; PRIOR APPLICATION NUMBER: 60/025,204  
; PRIOR FILING DATE: August 1, 1996  
; PRIOR APPLICATION NUMBER: 08/903,471  
; PRIOR FILING DATE: July 30, 1997  
; PRIOR APPLICATION NUMBER: 60/025,478  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/903,556  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/025,217  
; PRIOR FILING DATE: August 22, 1996  
; PRIOR APPLICATION NUMBER: 08/937,142  
; PRIOR FILING DATE: September 23, 1997  
; PRIOR APPLICATION NUMBER: 60/026,598  
; PRIOR FILING DATE: September 24, 1996  
; PRIOR APPLICATION NUMBER: 08/960,746  
; PRIOR FILING DATE: October 29, 1997  
; PRIOR APPLICATION NUMBER: 60/030,144  
; PRIOR FILING DATE: October 30, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/755,524  
; PRIOR FILING DATE: November 22, 1996  
; PRIOR APPLICATION NUMBER: 60/007,495  
; PRIOR FILING DATE: November 22, 1995  
; PRIOR APPLICATION NUMBER: 09/021,031  
; PRIOR FILING DATE: February 10, 1998  
; PRIOR APPLICATION NUMBER: 60/039,325  
; PRIOR FILING DATE: February 13, 1997  
; PRIOR APPLICATION NUMBER: 09/035,172  
; PRIOR FILING DATE: March 4, 1998  
; PRIOR APPLICATION NUMBER: 60/040,431  
; PRIOR FILING DATE: March 5, 1997  
; PRIOR APPLICATION NUMBER: 09/041,894  
; PRIOR FILING DATE: March 12, 1998  
; PRIOR APPLICATION NUMBER: 60/040,199  
; PRIOR FILING DATE: March 14, 1997  
; PRIOR APPLICATION NUMBER: 09/050,817  
; PRIOR FILING DATE: March 30, 1998  
; PRIOR APPLICATION NUMBER: 60/043,792  
; PRIOR FILING DATE: April 11, 1997

; PRIOR APPLICATION NUMBER: 09/074,999  
; PRIOR FILING DATE: May 8, 1998  
; PRIOR APPLICATION NUMBER: 60/048,431  
; PRIOR FILING DATE: May 29, 1997  
; PRIOR APPLICATION NUMBER: 09/107,592  
; PRIOR FILING DATE: June 30, 1998  
; PRIOR APPLICATION NUMBER: 60/052,751  
; PRIOR FILING DATE: July 1, 1997  
; PRIOR APPLICATION NUMBER: 09/094,079  
; PRIOR FILING DATE: June 9, 1998  
; PRIOR APPLICATION NUMBER: 60/049,975  
; PRIOR FILING DATE: June 13, 1997  
; NUMBER OF SEQ ID NOS: 35654  
; SOFTWARE: PERL Program  
; SEQ ID NO 21427  
; LENGTH: 204  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte ID No: hu00323909  
US-09-340-2108-21427

Query Match 1.0%; Score 16; DB 5; Length 204;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 841 gcctttaattttcttg 856  
|||||  
Db 115 gcctttaattttcttg 130

Search completed: April 27, 2002, 07:43:52  
Job time: 17084 sec



---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 01:39:32 ; Search time 8172.53 seconds

(Without alignments)

4336.229 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638

Sequence: 1 atgtcgccaccagcaactc.....tgcttcgtctcctgctaa 1638

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Searched: 21979536 seqs, 10817449327 residues

Word size : 0

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_NA\_Main:\*

- 1: /cgn2\_6/ptodata/2/pna/US06\_PCTUS\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*
- 8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*
- 23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*
- 25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*
- 26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*
- 27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*
- 28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*
- 29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*
- 30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*
- 31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*
- 32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*
- 33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*
- 34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq.\*
- 35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq.\*
- 36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq.\*
- 37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq.\*
- 38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq.\*
- 39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*
- 40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*
- 41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*
- 42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*
- 43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*

- 44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*
- 45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*
- 46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*
- 47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*
- 48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*
- 49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*
- 50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*
- 51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*
- 52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*
- 53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*
- 54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*
- 55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*
- 56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*
- 57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*
- 58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*
- 59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*
- 60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*
- 61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*
- 62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*
- 63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*
- 64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*
- 65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*
- 66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*
- 67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*
- 68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*
- 69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*
- 70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*
- 71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*
- 72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*
- 73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*
- 74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*
- 75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1638	100.0	1638	17 US-09-351-224-2	Sequence 2, Appli
2	1638	100.0	1638	17 US-09-351-224E-2	Sequence 2, Appli
3	1638	100.0	1638	17 US-09-351-823-2	Sequence 2, Appli
4	1638	100.0	1638	26 US-09-677-488-2	Sequence 2, Appli
5	1638	100.0	1638	26 US-09-677-488A-2	Sequence 2, Appli
6	1638	100.0	1638	26 US-09-677-682-2	Sequence 2, Appli
7	1638	100.0	1638	26 US-09-677-682A-2	Sequence 2, Appli
8	1638	100.0	1638	26 US-09-677-682B-2	Sequence 2, Appli
9	1638	100.0	1638	33 US-09-882-694-2	Sequence 2, Appli
10	1638	100.0	1638	33 US-09-882-694A-2	Sequence 2, Appli
11	1184	72.3	1691	17 US-09-351-224-1	Sequence 1, Appli
12	1184	72.3	1691	17 US-09-351-823-1	Sequence 1, Appli
13	1184	72.3	1691	17 US-09-677-488-1	Sequence 1, Appli
14	1184	72.3	1691	26 US-09-677-488A-1	Sequence 1, Appli
15	1184	72.3	1691	26 US-09-677-682-1	Sequence 1, Appli
16	1184	72.3	1691	26 US-09-677-682A-1	Sequence 1, Appli
17	1184	72.3	1691	26 US-09-677-682B-1	Sequence 1, Appli
18	1184	72.3	1691	33 US-09-882-694-1	Sequence 1, Appli
19	1184	72.3	1691	33 US-09-882-694A-1	Sequence 1, Appli
20	1184	72.3	1691	33 US-09-703-708-8432	Sequence 8432, Ap
21	22	1.3	1425	55 US-60-164-320-8432	Sequence 8432, Ap
22	22	1.3	1425	57 US-60-183-791-8432	Sequence 718, App
23	22	1.3	18433	28 US-09-703-708-718	Sequence 718, App
24	22	1.3	18433	55 US-60-164-320-718	Sequence 718, App
25	22	1.3	18433	57 US-60-183-791-718	Sequence 23, Appl
26	22	1.3	2087	75 US-60-365-076-23	Sequence 23, Appl
27	21	1.3	2109	75 US-60-365-076-25	Sequence 25, Appl
28	21	1.3	2123	75 US-60-365-076-26	Sequence 26, Appl
29	21	1.3	3615	75 US-60-365-076-24	Sequence 24, Appl
30	21	1.3	675	51 US-60-126-265-315	Sequence 315, App
31	20	1.2			

```
c 32 20 1.2 6332 33 US-09-897-516-1311 Sequence 1311, Ap
c 33 20 1.2 6332 33 US-09-897-516-1312 Sequence 1312, Ap
c 34 20 1.2 6332 33 US-09-897-516-1313 Sequence 1313, Ap
c 35 20 1.2 6332 33 US-09-897-516-1314 Sequence 1314, Ap
c 36 20 1.2 6332 33 US-09-897-516-1315 Sequence 1315, Ap
c 37 20 1.2 6332 33 US-09-897-516-1316 Sequence 1316, Ap
c 38 20 1.2 6332 60 US-60-215-161-1311 Sequence 1311, Ap
c 39 20 1.2 6332 60 US-60-215-161-1312 Sequence 1312, Ap
c 40 20 1.2 6332 60 US-60-215-161-1313 Sequence 1313, Ap
c 41 20 1.2 6332 60 US-60-215-161-1314 Sequence 1314, Ap
c 42 20 1.2 6332 60 US-60-215-161-1315 Sequence 1315, Ap
c 43 20 1.2 6332 60 US-60-215-161-1316 Sequence 1316, Ap
c 44 19 1.2 360 58 US-60-197-872-53324 Sequence 53324, A
c 45 19 1.2 391 17 US-09-304-517A-242711 Sequence 242711, A
```

## ALIGNMENTS

```
RESULT 1
US-09-351-224-2
: Sequence 2, Application US/09351224
: GENERAL INFORMATION:
: APPLICANT: Duick, Jon
: TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification
: FILE REFERENCE: 5718-111
: CURRENT APPLICATION NUMBER: US/09/351,224
: CURRENT FILING DATE: 1999-07-12
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1638
: TYPE: DNA
: ORGANISM: Exophiala spinifera
: FEATURE:
: OTHER INFORMATION: flavin monooxygenase, coding sequence
US-09-351-224-2
```

```
Query Match 100.0%; Score 1638; DB 17; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcgccaccagcaactccagagcgattgttccgtcgcatcgagcgccatcactggt 60
D 1 atgtcgccaccagcaactccagagcgattgttccgtcgcatcgagcgccatcactggt 60
QY 61 ggaagccgacctcagagcgatctctgtgtacaaattggaattcgaaagctcaactcaagcc 120
D 61 ggaagccgacctcagagcgatctctgtgtacaaattggaattcgaaagctcaactcaagcc 120
QY 121 aaattctcgaggagcccccgattttggcgcgctctggcactggaacccgctaccctggc 180
D 121 aaattctcgaggagcccccgattttggcgcgctctggcactggaacccgctaccctggc 180
QY 181 gctcggtgtgattcgagagcgccctctaccactgaacattcccgaaagtatggaagac 240
D 181 gctcggtgtgattcgagagcgccctctaccactgaacattcccgaaagtatggaagac 240
QY 241 tgaagctgtctgacgctatctctgaccagaaagagttgtgtcatatgttccaccactgt 300
D 241 tgaagctgtctgacgctatctctgaccagaaagagttgtgtcatatgttccaccactgt 300
QY 301 gacaagatccggggttgagaaagacgtctactctcgagagctgaggtggtgtgagcgcg 360
D 301 gacaagatccggggttgagaaagacgtctactctcgagagctgaggtggtgtgagcgcg 360
QY 361 tatgccagagatctgggacacctggactgtcaagacgtcgactgtgacgagga 420
D 361 tatgccagagatctgggacacctggactgtcaagacgtcgactgtgacgagga 420
QY 421 aagtatctattctgcgtacgggtgtgtccaaaggagacactcccgcaactccccgc 480
D 421 aagtatctattctgcgtacgggtgtgtccaaaggagacactcccgcaactccccgc 480
```

```
Db 421 aagtatctattctgcgtacgggtgtgtccaaaggagacactcccgcaactccccgc 480
QY 481 ctgcccatttcaacggggaaggtgattcattctgagtgccctggcagcaagactctgacgca 540
D 481 ctgcccatttcaacggggaaggtgattcattctgagtgccctggcagcaagactctgacgca 540
QY 541 gagggccagagagtcgcccgtcattcgtgcccggggccacaaagcatccagattgttcaggag 600
D 541 gagggccagagagtcgcccgtcattcgtgcccggggccacaaagcatccagattgttcaggag 600
QY 601 ttgcccagaagactcaccaggttaaccattgtttatgcaaaagccgagattgttctgcc 660
D 601 ttgcccagaagactcaccaggttaaccattgtttatgcaaaagccgagattgttctgcc 660
QY 661 atgcggcaacgaacgatgtagtagaagcagacagcctggaaagccctactaccaccacg 720
D 661 atgcggcaacgaacgatgtagtagaagcagacagcctggaaagccctactaccaccacg 720
QY 721 ctgtttgaagcgagtcgaaagtcctcgatttgattcccggtccagggcaccgtcgattggc 780
D 721 ctgtttgaagcgagtcgaaagtcctcgatttgattcccggtccagggcaccgtcgattggc 780
QY 781 atctttgaagtcagcccgagcagcgagcgagcctatttcgaagagattgtgagcgtggg 840
D 781 atctttgaagtcagcccgagcagcgagcgagcctatttcgaagagattgtgagcgtggg 840
QY 841 gcttttaattttctgttgcagtagaccgagaagtcattggttgacaaagggccaaaccga 900
D 841 gcttttaattttctgttgcagtagaccgagaagtcattggttgacaaagggccaaaccga 900
QY 901 ctggctctgactctcgggcccacaaagactcgtatctcgtatcgtcaatcccgccaaagaga 960
D 901 ctggctctgactctcgggcccacaaagactcgtatctcgtatcgtcaatcccgccaaagaga 960
QY 961 gatctcagtcctctggagcgccgactggttcgatttcgataccaaagcctccccactgag 1020
D 961 gatctcagtcctctggagcgccgactggttcgatttcgataccaaagcctccccactgag 1020
QY 1021 agcgactactacgaataatgtggacaagccgagcgtcgaaattgtgaatctagaacaatcg 1080
D 1021 agcgactactacgaataatgtggacaagccgagcgtcgaaattgtgaatctagaacaatcg 1080
QY 1081 cccatttggtctgttacaagaacaggtgtgctcttgagtgacggcgagcaaggggaatgc 1140
D 1081 cccatttggtctgttacaagaacaggtgtgctcttgagtgacggcgagcaaggggaatgc 1140
QY 1141 gacacgactgctgagcgagcgttttcgaagatttcactgactcactgacacatatgggc 1200
D 1141 gacacgactgctgagcgagcgttttcgaagatttcactgactcactgacacatatgggc 1200
QY 1201 ttgaaaaacaaagcacgagtggaacctgaaaggaggtgtggaagatggtgcatatctactat 1260
D 1201 ttgaaaaacaaagcacgagtggaacctgaaaggaggtgtggaagatggtgcatatctactat 1260
QY 1261 atgggagctctctctcattggcttcccccaatgctctctctcgtcgccacgggtcgaagcccg 1320
D 1261 atgggagctctctctcattggcttcccccaatgctctctctcgtcgccacgggtcgaagcccg 1320
QY 1321 accgtctctcccaagcccaacgatacattagaaacccagtcgactgtgactgcgcgataca 1380
D 1321 accgtctctcccaagcccaacgatacattagaaacccagtcgactgtgactgcgcgataca 1380
QY 1381 atygaaagtgtgagggccgagcacgcccacgtccgtttgaggagcaagaaatcagcaagag 1440
D 1381 atygaaagtgtgagggccgagcacgcccacgtccgtttgaggagcaagaaatcagcaagag 1440
QY 1441 gcatggttcgattatgatttgcgaagatgaaacagacactctgttcccccttgacggattcg 1500
D 1441 gcatggttcgattatgatttgcgaagatgaaacagacactctgttcccccttgacggattcg 1500
QY 1501 tgggtgagctggaagcaacatccctgggaaagcaacacactgtttaaccttcattagcgagg 1560
D 1501 tgggtgagctggaagcaacatccctgggaaagcaacacactgtttaaccttcattagcgagg 1560
```

```
QY 1561 attgctctctatgagcagatctgtcaagagaaggttgcccaattggatgggtttgatctg 1620
|||||
Db 1561 attgctctctatgagcagatctgtcaagagaaggttgcccaattggatgggtttgatctg 1620

QY 1621 ctctatgctcctgctctaa 1638
|||||
Db 1621 ctctatgctcctgctctaa 1638

RESULT 2
US-09-351-224E-2
; Sequence 2, Application US/09351224E
; GENERAL INFORMATION:
; APPLICANT: Duivick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 5718-111
; CURRENT APPLICATION NUMBER: US/09/351,224E
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: flavin monooxygenase, fully spliced
US-09-351-224E-2

Query Match 100.0%; Score 1638; DB 17; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtggccacagcaactcagaggcgattgttcgtgcgcatgcgacgcaactatcatt 60
Db 1 atgtggccacagcaactcagaggcgattgttcgtgcgcatgcgacgcaactatcatt 60

QY 61 ggagccgctcagcgccatctctgtgtacaaaattgcgaagctcagactcaacgcc 120
Db 61 ggagccgctcagcgccatctctgtgtacaaaattgcgaagctcagactcaacgcc 120

QY 121 aaatcttcgaggagcccccatttggcgcgctctggcactggagaccgtaccctggc 180
Db 121 aaatcttcgaggagcccccatttggcgcgctctggcactggagaccgtaccctggc 180

QY 181 gtcgtgtgtatcgagagccctctacaaactgaacattcccaagtatggaaagac 240
Db 181 gtcgtgtgtatcgagagccctctacaaactgaacattcccaagtatggaaagac 240

QY 241 tggacctgtctgcgctatcttcgaccagaaagattgtgtcatatttaccactgt 300
Db 241 tggacctgtctgcgctatcttcgaccagaaagattgtgtcatatttaccactgt 300

QY 301 gacaagatccgggcttgagaaagacgtctacttcgagactgaggtgtgatgcgag 360
Db 301 gacaagatccgggcttgagaaagacgtctacttcgagactgaggtgtgatgcgag 360

QY 361 tatgcagagatctggcactggactgtcaagacgtcggtggccatgttcgcagcgca 420
Db 361 tatgcagagatctggcactggactgtcaagacgtcggtggccatgttcgcagcgca 420

QY 421 aagatctctatctcgtacggggtgtgtccacaggaagcaactcccgactcccgccg 480
Db 421 aagatctctatctcgtacggggtgtgtccacaggaagcaactcccgactcccgccg 480
```

```
QY 481 ctgcgcgatttcaacgggaaggtgatcattcattcagtgctcctgggcacgaagacttcgacgca 540
|||||
Db 481 ctgcgcgatttcaacgggaaggtgatcattcattcagtgctcctgggcacgaagacttcgacgca 540

QY 541 gagggccagagagtcgcgtcattcgttcacggggccacaagaatccagattgttcaagag 600
|||||
Db 541 gagggccagagagtcgcgtcattcgttcacggggccacaagaatccagattgttcaagag 600

QY 601 ttggccaagaaggtgacacaggttaacattgtttatgcgaagggcgagctatgtctgccc 660
|||||
Db 601 ttggccaagaaggtgacacaggttaacattgtttatgcgaagggcgagctatgtctgccc 660

QY 661 atgcggcaacgacgatgtagaagcagacagacagcctgggaagggcctactaccacag 720
|||||
Db 661 atgcggcaacgacgatgtagaagcagacagacagcctgggaagggcctactaccacag 720

QY 721 ctggttgaaagcagatcgaagtctcgaattcgaattcccggtccagacacgtcggtggc 780
|||||
Db 721 ctggttgaaagcagatcgaagtctcgaattcgaattcccggtccagacacgtcggtggc 780

QY 781 atcttgaagtgcagccccgagcagcggggagcctatttcgaagagttgtgggagcgtggg 840
|||||
Db 781 atcttgaagtgcagccccgagcagcggggagcctatttcgaagagttgtgggagcgtggg 840

QY 841 gctttaattttctgtgctgcagtcaccgagaagtcattggttgacaaaaggcccaaccga 900
|||||
Db 841 gctttaattttctgtgctgcagtcaccgagaagtcattggttgacaaaaggcccaaccga 900

QY 901 ctggtctatgactttctggggccaaaagactcgtatctcgtatcgttcactcgcgcgaagaa 960
|||||
Db 901 ctggtctatgactttctggggccaaaagactcgtatctcgtatcgttcactcgcgcgaagaa 960

QY 961 gatctcagtgctcctctggagcccgctactggttcggttaccgaagggcctccactggag 1020
Db 961 gatctcagtgctcctctggagcccgctactggttcggttaccgaagggcctccactggag 1020

QY 1021 agcgactactacgaattctggacaagccgagcgtcgaattgtgaaatctagacaatcg 1080
Db 1021 agcgactactacgaattctggacaagccgagcgtcgaattgtgaaatctagacaatcg 1080

QY 1081 cccattgtgctgttacaagacaggtgtgctcttgaatgaacgagcagcagcagagggatcg 1140
|||||
Db 1081 cccattgtgctgttacaagacaggtgtgctcttgaatgaacgagcagcagcagagggatcg 1140

QY 1141 gacacgatcgtctggcgacgggtttcgacagtttccactggctccatgtgacacatatgggc 1200
Db 1141 gacacgatcgtctggcgacgggtttcgacagtttccactggctccatgtgacacatatgggc 1200

QY 1201 ttgaaaaacaagcagcggatggacctgaaggaggtgtggaaagatggcatatctactat 1260
Db 1201 ttgaaaaacaagcagcggatggacctgaaggaggtgtggaaagatggcatatctactat 1260

QY 1261 atggagatctctctcattggttcccaatgccttcttcgtccacgctcaagcccg 1320
Db 1261 atggagatctctctcattggttcccaatgccttcttcgtccacgctcaagcccg 1320

QY 1321 accgtcttccaaagcccaacgacatcatagaacccaagtcgactgtatgcgcgataca 1380
Db 1321 accgtcttccaaagcccaacgacatcatagaacccaagtcgactgtatgcgcgataca 1380

QY 1381 attgcaagtttgagggccgagcacccacgtccgttgaggcgacgaatcagcacaaag 1440
Db 1381 attgcaagtttgagggccgagcacccacgtccgttgaggcgacgaatcagcacaaag 1440

QY 1441 gcatgctatctatgatgccaagatgaacgacacactctgttccctctgacggattcg 1500
|||||
Db 1441 gcatgctatctatgatgccaagatgaacgacacactctgttccctctgacggattcg 1500

QY 1501 tggtagctggaggcaacatccctgggaaagcaacagctgtttaaoccttcatagcgagg 1560
|||||
Db 1501 tggtagctggaggcaacatccctgggaaagcaacagctgtttaaoccttcatagcgagg 1560

QY 1561 attgctctctatgagcagatctgtcaagagaaggtggccaattggggtttgatctg 1620
```

```
|||||
Db 1561 attgctctctatgagcagactctgcaagagaagggtggccaattgggattggattgagtgtg 1620
QY 1621 cttcatgctccctgcttaa 1638
|||||
Db 1621 cttcatgctccctgcttaa 1638

RESULT
3
US-09-351-823-2
; Sequence 2, Application US/09351823
; GENERAL INFORMATION:
; APPLICANT: Duwick, Jon
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification
; FILE REFERENCE: 5718-56
; CURRENT APPLICATION NUMBER: US/09/351,823
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; OTHER INFORMATION: flavin monooxygenase, coding sequence
US-09-351-823-2

Query Match 100.0%; Score 1638; DB 17; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcggccaccagcaactccagagcgattgttcgctgcgcatcgacgcgcacatcattgtt 60
Db 1 atgtcggccaccagcaactccagagcgattgttcgctgcgcatcgacgcgcacatcattgtt 60

QY 61 ggaacggcctcagcgccatctctgtgtgtacaaattgcgaagctcagactcaacgccc 120
Db 61 ggaacggcctcagcgccatctctgtgtgtacaaattgcgaagctcagactcaacgccc 120

QY 121 aaaaattcggaggagcccccgattttggcgcgctctggcactggaaacgctacccctggc 180
Db 121 aaaaattcggaggagcccccgattttggcgcgctctggcactggaaacgctacccctggc 180

QY 181 gctcgtgttgatcggagagccctctaccacactgaacattcccgaaagtatgaaagac 240
Db 181 gctcgtgttgatcggagagccctctaccacactgaacattcccgaaagtatgaaagac 240

QY 241 tggacctggcttgccgctatctgacccagaaagagttgctgtcatatgttccaccactgt 300
Db 241 tggacctggcttgccgctatctgacccagaaagagttgctgtcatatgttccaccactgt 300

QY 301 gacaagatccggggcttgagaaaagacgtctactctcgagctgaggtggttgtagtcgcg 360
Db 301 gacaagatccggggcttgagaaaagacgtctactctcgagctgaggtggttgtagtcgcg 360

QY 361 tatccagagatctggcaccctgagactgtcaagacgtcgctgcccattgttgagagcgca 420
Db 361 tatccagagatctggcaccctgagactgtcaagacgtcgctgcccattgttgagagcgca 420

QY 421 aagtatctatctcgtcaggggtgtctccacagaaagacactcccgacactcccgcc 480
Db 421 aagtatctatctcgtcaggggtgtctccacagaaagacactcccgacactcccgcc 480

QY 481 ctgcgcgatttcacggggaagtgtattcattcagctgctggcagcaagactcttcgacgca 540
Db 481 ctgcgcgatttcacggggaagtgtattcattcagctgctggcagcaagactcttcgacgca 540

QY 541 gagggccagagagtcgcccctcagctcagtcggcgccacacacacagatgtttcagag 600
Db 541 gagggccagagagtcgcccctcagctcagtcggcgccacacacacagatgtttcagag 600

QY 601 ttggccaagaagcgtgaccaggttaacattgtttatgcaagaagccgagctattgtctgccc 660
```

```
|||||
Db 601 ttggccaagaaggctgaccaggtaaccattgtttatgcgaaggccgagctattgtctgcgcc 660
QY 661 atgaggcaacgaacgatggtagaagcaacagacagccttggaagcctactaccaccacg 720
Db 661 atgaggcaacgaacgatggtagaagcaacagacagccttggaagcctactaccaccacg 720

QY 721 ctgtttgaagcagtcgaaagctcggatttggatttcccgttccaggtccaggtcgtgtggtg 780
Db 721 ctgtttgaagcagtcgaaagctcggatttggatttcccgttccaggtcgtgtggtg 780

QY 781 atotttgaagtcagcccccagcagcagcgggagcctatttccgaagagttgtgggagcgtggg 840
Db 781 atotttgaagtcagcccccagcagcagcgggagcctatttccgaagagttgtgggagcgtggg 840

QY 841 gcttttaattttctgttccagtaaccgagaagtcattgttgacaaaagggcccaaccga 900
Db 841 gcttttaattttctgttccagtaaccgagaagtcattgttgacaaaagggcccaaccga 900

QY 901 ctggtctatgacttctgggccccaaagactcgtatctgctatcgtcaatccggcaagaga 960
Db 901 ctggtctatgacttctgggccccaaagactcgtatctgctatcgtcaatccggcaagaga 960

QY 961 gatctcatgctctctggagccgcgctactgttccgttaccagagcgtcccccactggag 1020
Db 961 gatctcatgctctctggagccgcgctactgttccgttaccagagcgtcccccactggag 1020

QY 1021 agcgactactacgaattgctggcaagccgagcgttcgaattgtgaatctagaacaaatcg 1080
Db 1021 agcgactactacgaattgctggcaagccgagcgttcgaattgtgaatctagaacaaatcg 1080

QY 1081 cccatttggctgtgtacaaagacagatgtgactcttgagtgagcggcagcagaggaatgc 1140
Db 1081 cccatttggctgtgtacaaagacagatgtgactcttgagtgagcggcagcagaggaatgc 1140

QY 1141 gacacgatcgttctggcgacgggttttcgacagtttcaacttgcacttgcacacatatgggc 1200
Db 1141 gacacgatcgttctggcgacgggttttcgacagtttcaacttgcacttgcacacatatgggc 1200

QY 1201 ttgaaaacaaagcagcagtgaggcctgaaggaggtgtggaagatgagcatatctacttat 1260
Db 1201 ttgaaaacaaagcagcagtgaggcctgaaggaggtgtggaagatgagcatatctacttat 1260

QY 1261 atggagctcttctctctatggcttcccccaatggccttcttctgctgcgcacgtcgaagcccg 1320
Db 1261 atggagctcttctctctatggcttcccccaatggccttcttctgctgcgcacgtcgaagcccg 1320

QY 1321 accgtctcttccaaagcccaacgataatagaacccaagtcgacttgcacacatacaca 1380
Db 1321 accgtctcttccaaagcccaacgataatagaacccaagtcgacttgcacacatacaca 1380

QY 1381 attgcaagttggaggccgagcgcacgctccgttggaggcagcgaataatcagcaacagag 1440
Db 1381 attgcaagttggaggccgagcgcacgctccgttggaggcagcgaataatcagcaacagag 1440

QY 1441 gcatggtcgattatgattgccaaagatgaacgagacacactcttcccttgcagcagattcg 1500
Db 1441 gcatggtcgattatgattgccaaagatgaacgagacacactcttcccttgcagcagattcg 1500

QY 1501 tgggtggactggaggcaacataccctgggaaagaaacagctgctttaaactctataggcggg 1560
Db 1501 tgggtggactggaggcaacataccctgggaaagaaacagctgctttaaactctataggcggg 1560

QY 1561 attgctctctatgagcagatctgtcgaagagaaggttggccaattgggattggattgtg 1620
Db 1561 attgctctctatgagcagatctgtcgaagagaaggttggccaattgggattggattgtg 1620

QY 1621 cttcatgctccctgcttaa 1638
Db 1621 cttcatgctccctgcttaa 1638
```

RESULT 4

US-09-677-488-2

; Sequence 2, Application US/09677488

; GENERAL INFORMATION:

; APPLICANT: Duivick, Jon

; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification

; FILE REFERENCE: 35718/204100 (5718-111A)

; CURRENT APPLICATION NUMBER: US/09/677,488

; CURRENT FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: 09/351,224

; PRIOR FILING DATE: 1999-07-12

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1638

; TYPE: DNA

; ORGANISM: *Exophiala spinifera*

; FEATURE:

; OTHER INFORMATION: flavin monooxygenase, fully spliced cDNA

US-09-677-488-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	atgtcgccacagcaactccagagggcgattgttcgctgcagatgcagcgccatcatcgtt	60
Db	1	atgtcgccacagcaactccagagggcgattgttcgctgcagatgcagcgccatcatcgtt	60
Qy	61	ggagcgccgctcagcgcatctctgctgtgtacaaaattgcgaagctcagactcaacgcc	120
Db	61	ggagcgccgctcagcgcatctctgctgtgtacaaaattgcgaagctcagactcaacgcc	120
Qy	121	aaaattctcagggagcccgattttggcggtcttggcactggaacgcgtaccctggc	180
Db	121	aaaattctcagggagcccgattttggcggtcttggcactggaacgcgtaccctggc	180
Qy	181	gctcgtgtgattcgagagccctctaccgaactgaacattcccggaagtgtgaagac	240
Db	181	gctcgtgtgattcgagagccctctaccgaactgaacattcccggaagtgtgaagac	240
Qy	241	tggaactgtcttcgcgtatctccgacagaaagattgctgctcatatttcaactgtt	300
Db	241	tggaactgtcttcgcgtatctccgacagaaagattgctgctcatatttcaactgtt	300
Qy	301	gacaaatccgggcttgagaaagacgtctacttcggagctgaggtgttgatgcggg	360
Db	301	gacaaatccgggcttgagaaagacgtctacttcggagctgaggtgttgatgcggg	360
Qy	361	tatgcagagatctggcacctggactgtcaagacgtcgctggccatgttgcgacggca	420
Db	361	tatgcagagatctggcacctggactgtcaagacgtcgctggccatgttgcgacggca	420
Qy	421	aagtatctatctcgtacaggggttctccacagaaacacactccgcactcccgcc	480
Db	421	aagtatctatctcgtacaggggttctccacagaaacacactccgcactcccgcc	480
Qy	481	ctcgccgatttaacgggaaggtgatttatttcagtgctcgtggcgaagacttcgaagca	540
Db	481	ctcgccgatttaacgggaaggtgatttatttcagtgctcgtggcgaagacttcgaagca	540
Qy	541	gagggccagagatcgccgtcatcgttgcgggggcccacaaagcatccagattgttcagg	600
Db	541	gagggccagagatcgccgtcatcgttgcgggggcccacaaagcatccagattgttcagg	600
Qy	601	ttggccaagaaggtgacaggttaacctgtttatgcgaaggccgagctattgtctgcc	660
Db	601	ttggccaagaaggtgacaggttaacctgtttatgcgaaggccgagctattgtctgcc	660
Qy	661	atcgccgaacagcatgattaggaacagacagacgcttggagccttactaccctacg	720
Db	661	atcgccgaacagcatgattaggaacagacagacgcttggagccttactaccctacg	720

RESULT 5

US-09-677-488A-2

; Sequence 2, Application US/09677488A

; GENERAL INFORMATION:

; APPLICANT: Duivick, Jon

; APPLICANT: Maddox, Joyce

; APPLICANT: Gilliam, Jacob

; APPLICANT: Folkerts, Otto

APPLICANT: Crasta, Oswald R.  
TITLE OF INVENTION: Compositions and Methods for Fumonisins  
FILE REFERENCE: 35718/204100  
CURRENT APPLICATION NUMBER: US/09/677,488A  
CURRENT FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 09/351,224  
PRIOR FILING DATE: 1999-07-12  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1638  
TYPE: DNA  
ORGANISM: Exophiala spinifera  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)-(10)  
OTHER INFORMATION: flavin monooxygenase, fully spliced  
US-09-677-488A-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcggccacagcaactccagagggattgttcctgcgcatgagcgcacatcatcgtt 60  
DB 1 atgtcggccacagcaactccagagggattgttcctgcgcatgagcgcacatcatcgtt 60  
QY 61 gtagcggccctcagcggcatctctgtgtatcaaaattgcgaagctcagactcaacgc 120  
DB 61 gtagcggccctcagcggcatctctgtgtatcaaaattgcgaagctcagactcaacgc 120  
QY 121 aaaaatttcggaggagcccccagatttggcggtctgtgcactggaacggctaccctggc 180  
DB 121 aaaaatttcggaggagcccccagatttggcggtctgtgcactggaacggctaccctggc 180  
QY 181 gtcgtgtgtattcggagagcccttctacaaactgaaacatccccaagatgtggaagac 240  
DB 181 gtcgtgtgtattcggagagcccttctacaaactgaaacatccccaagatgtggaagac 240  
QY 241 tggactcgtgttcgcgtatcttcgacgaagaagattgtgtctatatttaccactgt 300  
DB 241 tggactcgtgttcgcgtatcttcgacgaagaagattgtgtctatatttaccactgt 300  
QY 301 gacaagatccgggcttgaaaaagacgtctacttcgagctgaggtggttgatgcagg 360  
DB 301 gacaagatccgggcttgaaaaagacgtctacttcgagctgaggtggttgatgcagg 360  
QY 361 tatgcagagatctggcaccctggactgtcaagacgtcggtcggccatgtttgcgacggca 420  
DB 361 tatgcagagatctggcaccctggactgtcaagacgtcggtcggccatgtttgcgacggca 420  
QY 421 aagtatctcattctcgtacgggttgctccacagaaagacacactcccgactcccgcc 480  
DB 421 aagtatctcattctcgtacgggttgctccacagaaagacacactcccgactcccgcc 480  
QY 481 ctgcgcgatttcaacgggaaggtgattcattcagtgctgctggcagaaagacttcagcga 540  
DB 481 ctgcgcgatttcaacgggaaggtgattcattcagtgctgctggcagaaagacttcagcga 540  
QY 541 gaggcagagagctcgccatcgatgcggggccacaaagcatccagattgttcaggag 600  
DB 541 gaggcagagagctcgccatcgatgcggggccacaaagcatccagattgttcaggag 600  
QY 601 ttggccagagaggtgaccaggtaaccatgtttatcgaaagccgagctattgtctgcc 660  
DB 601 ttggccagagaggtgaccaggtaaccatgtttatcgaaagccgagctattgtctgcc 660  
QY 661 atcgccgaacaaagcattgtagaagcaacagacagcgtggaaggcctactaccacag 720  
DB 661 atcgccgaacaaagcattgtagaagcaacagacagcgtggaaggcctactaccacag 720

QY 721 ctggtttgaagcagatcgaagctctcgatttgattcccggtccagacacgctgattgac 780  
DB 721 ctggtttgaagcagatcgaagctctcgatttgattcccggtccagacacgctgattgac 780  
QY 781 atctttgaagtcagccccgagcagcgggagccctatttcgaagagctgtgaggagcgtggg 840  
DB 781 atctttgaagtcagccccgagcagcgggagccctatttcgaagagctgtgaggagcgtggg 840  
QY 841 gcttttaatttcttctgttcgacagcagagtcattgttgacaaaagggcccaaccga 900  
DB 841 gcttttaatttcttctgttcgacagcagagtcattgttgacaaaagggcccaaccga 900  
QY 901 ctgggtctatgacttctgggcccacaaagactcgtatctcgttcaatccgccaagaga 960  
DB 901 ctgggtctatgacttctgggcccacaaagactcgtatctcgttcaatccgccaagaga 960  
QY 961 gatctcatggtctcttggagccgcgtactgttgcgttaccgaagcgtcccccactggag 1020  
DB 961 gatctcatggtctcttggagccgcgtactgttgcgttaccgaagcgtcccccactggag 1020  
QY 1021 agcgactactcgaatctgtggacagccgagcgtcgaaatttgtgaattctagaacaatcg 1080  
DB 1021 agcgactactcgaatctgtggacagccgagcgtcgaaatttgtgaattctagaacaatcg 1080  
QY 1081 cccattgtggctgtttacaagaagacaggtgtgtctcttgaagtcagcggcagcaagaggaatgc 1140  
DB 1081 cccattgtggctgtttacaagaagacaggtgtgtctcttgaagtcagcggcagcaagaggaatgc 1140  
QY 1141 gacacgacgtctgtcggcagcgggttttcgacagtttcaactggctcattgacacatatgggc 1200  
DB 1141 gacacgacgtctgtcggcagcgggttttcgacagtttcaactggctcattgacacatatgggc 1200  
QY 1201 ttgaaaaaacagcagcggagtggaacctgaaaggaggtgtggaagatggaatatactactat 1260  
DB 1201 ttgaaaaaacagcagcggagtggaacctgaaaggaggtgtggaagatggaatatactactat 1260  
QY 1261 atggagctctctctctatggtctccccaatgctcttcgctcgcacgctcaagcccg 1320  
DB 1261 atggagctctctctctatggtctccccaatgctcttcgctcgcacgctcaagcccg 1320  
QY 1321 accgtctcttccacggcccaacgacatagaaaccccaagtcgacttgcgcgacataca 1380  
DB 1321 accgtctcttccacggcccaacgacatagaaaccccaagtcgacttgcgcgacataca 1380  
QY 1381 attcaaaagttgagcggcagcgcagctccgtttgagcgcacgaaatcagacaagag 1440  
DB 1381 attcaaaagttgagcggcagcgcagctccgtttgagcgcacgaaatcagacaagag 1440  
QY 1441 gcatggtcgtatgattgataatgcaagatgaacgagacacactctgttcccttgacggattcg 1500  
DB 1441 gcatggtcgtatgattgataatgcaagatgaacgagacacactctgttcccttgacggattcg 1500  
QY 1501 tggtagactggaggcaacatccctgggaagcaacacgctgtcttaaccttataggcggg 1560  
DB 1501 tggtagactggaggcaacatccctgggaagcaacacgctgtcttaaccttataggcggg 1560  
QY 1561 attgctctctatgacagatctgttcaagagaaagtggtcccaattggggtttgattgtg 1620  
DB 1561 attgctctctatgacagatctgttcaagagaaagtggtcccaattggggtttgattgtg 1620  
QY 1621 ctctactgctccctactaa 1638  
DB 1621 ctctactgctccctactaa 1638

RESULT 6

US-09-677-682-2  
; Sequence 2, Application US/09677682  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fumonisins Detoxification  
; FILE REFERENCE: 35718/204101 (5718-111B)  
; CURRENT APPLICATION NUMBER: US/09/677,682



; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: flavin monooxygenase, fully spliced cDNA  
us-09-677-682-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atctcgccaccagcaactccagagcgattgtccgtcgcatgagcgccatcatcggtt 60  
Db 1 atctcgccaccagcaactccagagcgattgtccgtcgcatgagcgccatcatcggtt 60

Qy 61 gggagccgctcagcggaatctctgtgtgtacaaattcgaaagctcagactcaagcc 120  
Db 61 gggagccgctcagcggaatctctgtgtgtacaaattcgaaagctcagactcaagcc 120

Qy 121 aaatctcagggggcccccgtatttggcggtctggactggaaacgctaccctggc 180  
Db 121 aaatctcagggggcccccgtatttggcggtctggactggaaacgctaccctggc 180

Qy 181 gctcgtgtgattcgagagccctcttaccactgaacattcccgaaagtatgaaagac 240  
Db 181 gctcgtgtgattcgagagccctcttaccactgaacattcccgaaagtatgaaagac 240

Qy 241 tggactgtcttgcgtatcctgaccagaaagattgtctgatatgttcacactgt 300  
Db 241 tggactgtcttgcgtatcctgaccagaaagattgtctgatatgttcacactgt 300

Qy 301 gacagatccgggcttgagaaagacgtctacttcggagctgaggtgttgatgcgcg 360  
Db 301 gacagatccgggcttgagaaagacgtctacttcggagctgaggtgttgatgcgcg 360

Qy 361 tatgcagagatctgggacactgactgtcaagactcggtcgccatgttgagagcca 420  
Db 361 tatgcagagatctgggacactgactgtcaagactcggtcgccatgttgagagcca 420

Qy 421 aagtatctattctcgtacggggttgcctccaggaagacacactcccgactcccgcc 480  
Db 421 aagtatctattctcgtacggggttgcctccaggaagacacactcccgactcccgcc 480

Qy 481 ctgcgatttcaacgggaaggtgattcattctcagtgctggcgacgaagacttcgacga 540  
Db 481 ctgcgatttcaacgggaaggtgattcattctcagtgctggcgacgaagacttcgacga 540

Qy 541 gaggccagagatcgctcgtatcggtgcccggggccacagcatccagattgttcaggag 600  
Db 541 gaggccagagatcgctcgtatcggtgcccggggccacagcatccagattgttcaggag 600

Qy 601 ttggccaaagagctgacaggttaaccatttttatcgaaagccagatattgtctccc 660  
Db 601 ttggccaaagagctgacaggttaaccatttttatcgaaagccagatattgtctccc 660

Qy 661 atcgccacagacgatggataggaacagacagacagctggaagccactaccacag 720  
Db 661 atcgccacagacgatggataggaacagacagacagctggaagccactaccacag 720

Qy 721 ctgttgaagcagtgcaaatctcgtgatttcccggtccagggccagctcggttggc 780  
Db 721 ctgttgaagcagtgcaaatctcgtgatttcccggtccagggccagctcggttggc 780

Qy 781 atcttgaagtcagcccgagcagcgaggagcctatttcgaagattgttgagcgctggg 840  
Db 781 atcttgaagtcagcccgagcagcgaggagcctatttcgaagattgttgagcgctggg 840

Qy 841 gaccttaattttcttctgttccagttaccgagagtcattggtgacaaaaggccaaccca 900  
Db 841 gaccttaattttcttctgttccagttaccgagagtcattggtgacaaaaggccaaccca 900

Qy 901 ctggtctatgactttctggccaaaagactcgtatcgtatcgttcactccgggcaagaga 960  
Db 901 ctggtctatgactttctggccaaaagactcgtatcgtatcgttcactccgggcaagaga 960

Qy 961 gatctcattggtcctctgggagcccgctactcgtgttcgttaccagcgctccactggag 1020  
Db 961 gatctcattggtcctctgggagcccgctactcgtgttcgttaccagcgctccactggag 1020

Qy 1021 agcgactactcgaatctgggacagccgagcgtcgaaattgtgaattctagacaactcg 1080  
Db 1021 agcgactactcgaatctgggacagccgagcgtcgaaattgtgaattctagacaactcg 1080

Qy 1081 cccattgtggtctgtacaaagacaggtgtgctctgaatgtgaattctagacaactcg 1140  
Db 1081 cccattgtggtctgtacaaagacaggtgtgctctgaatgtgaattctagacaactcg 1140

Qy 1141 gacagatcgtgctgggagcggttttcgacagtttcaactgactcattcaacatatggc 1200  
Db 1141 gacagatcgtgctgggagcggttttcgacagtttcaactgactcattcaacatatggc 1200

Qy 1201 ttgaaaaaacagcgaggtggactgaaaggaggtgtggaagaatgagcattctacttat 1260  
Db 1201 ttgaaaaaacagcgaggtggactgaaaggaggtgtggaagaatgagcattctacttat 1260

Qy 1261 atggagtgcttctctcattggtcttcccaatgctcttcttctgcccagcggtcaggcccg 1320  
Db 1261 atggagtgcttctctcattggtcttcccaatgctcttcttctgcccagcggtcaggcccg 1320

Qy 1321 accgtcttcccaacggcccaacgacatagaaacccaagtcgacttgatcgccgataca 1380  
Db 1321 accgtcttcccaacggcccaacgacatagaaacccaagtcgacttgatcgccgataca 1380

Qy 1381 attgcaaatgtggagcccgagcgcacgtcgttgaggcgagcaaatcagcaagag 1440  
Db 1381 attgcaaatgtggagcccgagcgcacgtcgttgaggcgagcaaatcagcaagag 1440

Qy 1441 gcatggtcattatgattgccaagatgaacgagacacactctgtcccttgacagattcg 1500  
Db 1441 gcatggtcattatgattgccaagatgaacgagacacactctgtcccttgacagattcg 1500

Qy 1501 tggtagctgagggaacacatcccttgggaaagcaacacgtgtgttaaccttaagcgagg 1560  
Db 1501 tggtagctgagggaacacatcccttgggaaagcaacacgtgtgttaaccttaagcgagg 1560

Qy 1561 attgctctctatgagcagatctgcaagagaaggtggccaattgggattggttggatgtg 1620  
Db 1561 attgctctctatgagcagatctgcaagagaaggtggccaattgggattggttggatgtg 1620

Qy 1621 ettcattgctcctcgtctaa 1638  
Db 1621 ettcattgctcctcgtctaa 1638

RESULT 7  
US-09-677-682A-2  
; Sequence 2: Application US/09677682A  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 35718/204101  
; CURRENT APPLICATION NUMBER: US/09/677,682A  
; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224



; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: flavin monooxygenase, fully spliced  
US-09-677-682B-2

Query Match 100.0%; Score 1638; DB 26; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtcgccaccagcaactccagagcgattgttcgctgcgacgagccatcgtt 60  
Db 1 atgtcgccaccagcaactccagagcgattgttcgctgcgacgagccatcgtt 60

Qy 61 ggaagcgccctcagcgcatctctgtgtacaaaattcgaaagctcagactcaagcc 120  
Db 61 ggaagcgccctcagcgcatctctgtgtacaaaattcgaaagctcagactcaagcc 120

Qy 121 aaaaatttcgagggagcccgattttggcgcgctgtgcgactggaacgctaccctggc 180  
Db 121 aaaaatttcgagggagcccgattttggcgcgctgtgcgactggaacgctaccctggc 180

Qy 181 gctcgttgatttcgagagccctctaccactgaacattcccgaaagtatggaagac 240  
Db 181 gctcgttgatttcgagagccctctaccactgaacattcccgaaagtatggaagac 240

Qy 241 tggacgtctctgcgctacccctgaccagaaagattgtctgtcatattgtccaccctgt 300  
Db 241 tggacgtctctgcgctacccctgaccagaaagattgtctgtcatattgtccaccctgt 300

Qy 301 gacaagatccgggcttgagaaagcgtctacttcggagctgaggtgttgatgcgcg 360  
Db 301 gacaagatccgggcttgagaaagcgtctacttcggagctgaggtgttgatgcgcg 360

Qy 361 tatgcagagatctgggcaactggactgtcaagaactcggtcgccatgttgacagagca 420  
Db 361 tatgcagagatctgggcaactggactgtcaagaactcggtcgccatgttgacagagca 420

Qy 421 aagtatctattctcgtacgggtgtgtctccacagaaagcacactcccgactcccgcc 480  
Db 421 aagtatctattctcgtacgggtgtgtctccacagaaagcacactcccgactcccgcc 480

Qy 481 ctgcgcgatttcaacgggaaggtgattcattcgtgctgtgcgacgagacttcgacgca 540  
Db 481 ctgcgcgatttcaacgggaaggtgattcattcgtgctgtgcgacgagacttcgacgca 540

Qy 541 gaggccagagatccgctcgtatcggtgcggggccacaaagctccagatgttcaagag 600  
Db 541 gaggccagagatccgctcgtatcggtgcggggccacaaagctccagatgttcaagag 600

Qy 601 ttggccaaaggtgacaggttaaccatgtttatgcgaagggccagatattgtctgcc 660  
Db 601 ttggccaaaggtgacaggttaaccatgtttatgcgaagggccagatattgtctgcc 660

Qy 661 atcgcccaacgacatgtagtaggaacagacagacgctggaagcctactaccacag 720  
Db 661 atcgcccaacgacatgtagtaggaacagacagacgctggaagcctactaccacag 720

Qy 721 ctgttgagagcagtcgaaagctcgtgattggattcccggtccagaccgctcggtggc 780  
Db 721 ctgttgagagcagtcgaaagctcgtgattggattcccggtccagaccgctcggtggc 780

Qy 781 atctttgaagtcagcccgacgacgagggagccctatttcgaagagttgtggagcgtggg 840  
Db 781 atctttgaagtcagcccgacgacgagggagggcctatttcgaagagttgtggagcgtggg 840

Qy 841 gcctttaatttcttctgtgcccagtagccagagaaagtcatggttgacaaaaagggccaacgca 900  
Db 841 gcctttaatttcttctgtgcccagtagccagagaaagtcatggttgacaaaaagggccaacgca 900

Qy 901 ctggtctatgacttctgggcccacaaagactcgtatctcgtatcgtcaatcccggaagaga 960  
Db 901 ctggtctatgacttctgggcccacaaagactcgtatctcgtatcgtcaatcccggaagaga 960

Qy 961 gatctcatgctcctctgagcccgctactggttctggtaccacagcgtccccactggag 1020  
Db 961 gatctcatgctcctctgagcccgctactggttctggtaccacagcgtccccactggag 1020

Qy 1021 agcgactactacgaatgctgacaaagccgagcgtcgaatctgtaattctagaacaatcg 1080  
Db 1021 agcgactactacgaatgctgacaaagccgagcgtcgaatctgtaattctagaacaatcg 1080

Qy 1081 cccattgtgctgttacaagacaggggtgtctctttagtgacgagcaagaggaatgc 1140  
Db 1081 cccattgtgctgttacaagacaggggtgtctctttagtgacgagcaagaggaatgc 1140

Qy 1141 gacacgatcgtctgacgaggggttctcgacagtttctcactggctcattgacacatatgggc 1200  
Db 1141 gacacgatcgtctgacgaggggttctcgacagtttctcactggctcattgacacatatgggc 1200

Qy 1201 ttgaaaaacaacgacgaggtgacactgaagaggtgtggaagatgagcatatctacttat 1260  
Db 1201 ttgaaaaacaacgacgaggtgacactgaagaggtgtggaagatgagcatatctacttat 1260

Qy 1261 atggaggtctctctcgtggttctcccaatgctcttcttctgctgcacggctcaagcccg 1320  
Db 1261 atggaggtctctctcgtggttctcccaatgctcttcttctgctgcacggctcaagcccg 1320

Qy 1321 accgtctcttccaaagcccaacgatcatagaaacccaagctgacttgatcccgataca 1380  
Db 1321 accgtctcttccaaagcccaacgatcatagaaacccaagctgacttgatcccgataca 1380

Qy 1381 attgcaaaagtgtgagggccgacgacgacgttccgttgagggcagcaaatcagcaagag 1440  
Db 1381 attgcaaaagtgtgagggccgacgacgacgttccgttgagggcagcaaatcagcaagag 1440

Qy 1441 gcatgctcatatgatttcccaagatgaagagacacactctgtcccttgacgattcg 1500  
Db 1441 gcatgctcatatgatttcccaagatgaagagacacactctgtcccttgacgattcg 1500

Qy 1501 tggtagactggagggcaacatccctgggaaagcaacacgcttcttaaccttcatagggcg 1560  
Db 1501 tggtagactggagggcaacatccctgggaaagcaacacgcttcttaaccttcatagggcg 1560

Qy 1561 attgctctctatgagcagatctgtcaagagaaggtggccaattgggatgggttggatgtg 1620  
Db 1561 attgctctctatgagcagatctgtcaagagaaggtggccaattgggatgggttggatgtg 1620

Qy 1621 ctccaagctccctgctaa 1638  
Db 1621 ctccaagctccctgctaa 1638

RESULT 9  
US-09-882-694-2  
; Sequence 2, Application US/09882694  
; GENERAL INFORMATION:  
; APPLICANT: DuVick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 35718/208255  
; CURRENT APPLICATION NUMBER: US/09/882,694  
; CURRENT FILING DATE: 2001-06-12  
; PRIOR APPLICATION NUMBER: 09/351,224

```

; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 2
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Flavin monooxygenase, fully spliced
US-09-882-694-2

Query Match      100.0%; Score 1638; DB 33; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcggccacacgaactccagagcgattgttcgctgcgatgagcgccatcatcgtt 60
DB 1 atgtcggccacacgaactccagagcgattgttcgctgcgatgagcgccatcatcgtt 60

QY 61 gtagcggccctcagcggcattctgtgtacaaattgcgaagctcagactcaagccc 120
DB 61 gtagcggccctcagcggcattctgtgtacaaattgcgaagctcagactcaagccc 120

QY 121 aaaaattctcggaggagccccgattttggcggcgctgtgcactggaacgctaccctggc 180
DB 121 aaaaattctcggaggagccccgattttggcggcgctgtgcactggaacgctaccctggc 180

QY 181 gctcgtatgttcgagagcgcctctaccactgaacattcccgaaagtatgaagac 240
DB 181 gctcgtatgttcgagagcgcctctaccactgaacattcccgaaagtatgaagac 240

QY 241 tggacctggtcttcgctatctctgacccgaaagagtgctgtcatatgttcaacactgt 300
DB 241 tggacctggtcttcgctatctctgacccgaaagagtgctgtcatatgttcaacactgt 300

QY 301 gacaagatccggggcttgagaaagacgtctactctcggagctgagtggttgatgcgcg 360
DB 301 gacaagatccggggcttgagaaagacgtctactctcggagctgagtggttgatgcgcg 360

QY 361 tatgcagagatctgggacactgagctgtcaagagctgcgtggcattgttgacggca 420
DB 361 tatgcagagatctgggacactgagctgtcaagagctgcgtggcattgttgacggca 420

QY 421 aagtatctatctcgtacgggttgctccacagagacacactccgcgaactcccgcc 480
DB 421 aagtatctatctcgtacgggttgctccacagagacacactccgcgaactcccgcc 480

QY 481 ctgcgcgatttcaacgggaaggtgattcatctcagtgctggcagaaagatttcgacga 540
DB 481 ctgcgcgatttcaacgggaaggtgattcatctcagtgctggcagaaagatttcgacga 540

QY 541 gagggccagagatcgccgctcatcggtgcggggcccaagcatccagattgttcaggag 600
DB 541 gagggccagagatcgccgctcatcggtgcggggcccaagcatccagattgttcaggag 600

QY 601 ttggccaagaaggtgaccaggttaaccattgttatgcgaagccgagctattgtctgcc 660
DB 601 ttggccaagaaggtgaccaggttaaccattgttatgcgaagccgagctattgtctgcc 660

QY 661 atgcggcaacgaacgatggtagaagacgaacagacagcgtgggaagccctactaccacag 720
DB 661 atgcggcaacgaacgatggtagaagacgaacagacagcgtgggaagccctactaccacag 720

QY 721 ctgtttgaagcagctcgaagctcgtgatttgattcccggtccagggaccctcggttggc 780
DB 721 ctgtttgaagcagctcgaagctcgtgatttgattcccggtccagggaccctcggttggc 780

QY 781 atctttgaagtcagccccgagcggggagggctatttcgaagagttgtggagcgtggg 840
DB 781 atctttgaagtcagccccgagcggggagggctatttcgaagagttgtggagcgtggg 840

```

```

RESULT 10
US-09-882-694A-2
; Sequence 2: Application US/09882694A
; GENERAL INFORMATION:
; APPLICANT: Duwick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 5718-111
; CURRENT APPLICATION NUMBER: US/09/882.694A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 09/351,224

```

```
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Flavin monooxygenase, fully spliced
US-09-882-694A-2

Query Match      100.0%; Score 1638; DB 33; Length 1638;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtcgccaccagcaactccagagcgattgttcctgcgcgacgagccatcatcggtt 60
D 1 atgtcgccaccagcaactccagagcgattgttcctgcgcgacgagccatcatcggtt 60
QY 61 gtagcgccctcagcgacatctctgtgtacaaattgcgaagctcaagctcaagcgc 120
D 61 gtagcgccctcagcgacatctctgtgtacaaattgcgaagctcaagctcaagcgc 120
QY 121 aaattcttgaggagccccgattttggcggcgtctggaactggaaacccctaccctggc 180
D 121 aaattcttgaggagccccgattttggcggcgtctggaactggaaacccctaccctggc 180
QY 181 gtcgtgtgatttcggagagccctcttaccacgaacattcccggaagtatggaagac 240
D 181 gtcgtgtgatttcggagagccctcttaccacgaacattcccggaagtatggaagac 240
QY 241 tggacctgtcttcgacctatctgaccagaaagattgtctgtatgttaccactgt 300
D 241 tggacctgtcttcgacctatctgaccagaaagattgtctgtatgttaccactgt 300
QY 301 gacaagatccgggcttgaaagagcgtctacttcggagctgaggtggtgatgcgcgg 360
D 301 gacaagatccgggcttgaaagagcgtctacttcggagctgaggtggtgatgcgcgg 360
QY 361 tatgcagagatctggccacctggactgtcaagacgctcggctggccatgttggaagcga 420
D 361 tatgcagagatctggccacctggactgtcaagacgctcggctggccatgttggaagcga 420
QY 421 aagatctcattctcgtacggggttgctccacaggaagcaactcccgcaactcccgcc 480
D 421 aagatctcattctcgtacggggttgctccacaggaagcaactcccgcaactcccgcc 480
QY 481 ctgcgcgatttcaacgggaaggtgattcattcgtgctggcagcaagacttcgaagca 540
D 481 ctgcgcgatttcaacgggaaggtgattcattcgtgctggcagcaagacttcgaagca 540
QY 541 gaagccagagatgcgcgtcgtcgtcgttcgttcgttcgttcgttcgttcgttcgttc 600
D 541 gaagccagagatgcgcgtcgtcgtcgttcgttcgttcgttcgttcgttcgttcgttc 600
QY 601 ttggccaaagaagctgaccaggttaacattttatgcgaagcgagactattctccc 660
D 601 ttggccaaagaagctgaccaggttaacattttatgcgaagcgagactattctccc 660
QY 661 atgcggcaacgacgtgattgagaaacagacagacgctggaagcctactaccacag 720
D 661 atgcggcaacgacgtgattgagaaacagacagacgctggaagcctactaccacag 720
QY 721 ctgtttgaagcagctcgaagctcgtgattggtatcccggtccagggcaccgtcggtggc 780
D 721 ctgtttgaagcagctcgaagctcgtgattggtatcccggtccagggcaccgtcggtggc 780
QY 781 atcttgaagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 840
D 781 atcttgaagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 840
```

```
QY 841 gcccttaatttcttctgtccagtagcagagagtcattggttgacaaaaaggccaaacga 900
D 841 gcccttaatttcttctgtccagtagcagagagtcattggttgacaaaaaggccaaacga 900
QY 901 ctggtctatgactcttggtgccccaaagactcgtatcgtcatcgtcaatccggaagaga 960
D 901 ctggtctatgactcttggtgccccaaagactcgtatcgtcatcgtcaatccggaagaga 960
QY 961 gatctcattgctcctctggaaccccgactactggttcgtaccagcgctccctcactggag 1020
D 961 gatctcattgctcctctggaaccccgactactggttcgtaccagcgctccctcactggag 1020
QY 1021 aqcgactactcgaatgctggcaagccgagcgtcgaaaattgtgaattcagacaactcg 1080
D 1021 aqcgactactcgaatgctggcaagccgagcgtcgaaaattgtgaattcagacaactcg 1080
QY 1081 cccatttggtgctgttacaagagacagtgctcttgagtgacgagcagcaagagaggaatgc 1140
D 1081 cccatttggtgctgttacaagagacagtgctcttgagtgacgagcagcaagagaggaatgc 1140
QY 1141 gacacgactcgtctggcgacgggtttcgcacagtttcaactggctcattgacacatatgggc 1200
D 1141 gacacgactcgtctggcgacgggtttcgcacagtttcaactggctcattgacacatatgggc 1200
QY 1201 ttgaaaaaacaagcagcagtgagacctgaagggaggtgtgaaagatggcattacttat 1260
D 1201 ttgaaaaaacaagcagcagtgagacctgaagggaggtgtgaaagatggcattacttat 1260
QY 1261 atggagactctctctatggctcccaatgctctctctctctctctctctctctctctctct 1320
D 1261 atggagactctctctatggctcccaatgctctctctctctctctctctctctctctctct 1320
QY 1321 accgtctcttccaaaggcccaacgacatagaaaaacccaagtcgacttgatcgccgataca 1380
D 1321 accgtctcttccaaaggcccaacgacatagaaaaacccaagtcgacttgatcgccgataca 1380
QY 1381 attgcaaaagtgtggagccgacgacgcctccgtttgagggcagcagaaatcagcaagag 1440
D 1381 attgcaaaagtgtggagccgacgacgcctccgtttgagggcagcagaaatcagcaagag 1440
QY 1441 gcatggtcattgattgccaagatgaacgagcagacactctgttcccttgacggattcg 1500
D 1441 gcatggtcattgattgccaagatgaacgagcagacactctgttcccttgacggattcg 1500
QY 1501 tggtagactggaggaacacatccctgggaagcaacacgtgttacccttcattagcgcg 1560
D 1501 tggtagactggaggaacacatccctgggaagcaacacgtgttacccttcattagcgcg 1560
QY 1561 attgctctctatgagcagatgttcaagagaaggtggccaaattgggattgggtttgatgtg 1620
D 1561 attgctctctatgagcagatgttcaagagaaggtggccaaattgggattgggtttgatgtg 1620
QY 1621 cttcattgctccctgctaa 1638
D 1621 cttcattgctccctgctaa 1638
```

```
RESULT 11
US-09-351-224-1
; Sequence 1, Application US/09351224
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification
; FILE REFERENCE: 5718-111
; CURRENT APPLICATION NUMBER: US/09/351,224
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1691
; TYPE: DNA
; ORGANISM: Exophiala spinifera
```

; FEATURE:  
; OTHER INFORMATION: flavin monooxygenase with intron  
US-09-351-224-1

Query Match 72.3%; Score 1184; DB 17; Length 1691;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgtcgccacacagcaactcaagcgattgttcgctcgatcgacgacccatcatcgtt 60
DB 1 atgtcgccacacagcaactcaagcgattgttcgctcgatcgacgacccatcatcgtt 60
QY 61 ggaagcggcctcagcggaatctctgtgtgtacaaatgcgaagactcagactcaacgc 120
DB 61 ggaagcggcctcagcggaatctctgtgtgtacaaatgcgaagactcagactcaacgc 120
QY 121 aaattctcgaggggccccgattttggcggtgtgtgacactgaaacgctacccctggc 180
DB 121 aaattctcgaggggccccgattttggcggtgtgtgacactgaaacgctacccctggc 180
QY 181 gctcgtgtgattcgagagcgcctctctacaaactgaacattcccgaaagtatgaaagac 240
DB 181 gctcgtgtgattcgagagcgcctctctacaaactgaacattcccgaaagtatgaaagac 240
QY 241 tggacctgtgttcgctcattcctgacccagaagaggttgcgtcatatgttcacactgt 300
DB 241 tggacctgtgttcgctcattcctgacccagaagaggttgcgtcatatgttcacactgt 300
QY 301 gacagatccggggtctgagaaagacgctctctcggagctgaggtgggtgagcgcg 360
DB 301 gacagatccggggtctgagaaagacgctctctcggagctgaggtgggtgagcgcg 360
QY 361 tatgcagagatctgggcactgactgtcaagaactgctggtggtggtggtgagcgga 420
DB 361 tatgcagagatctgggcactgactgtcaagaactgctggtggtggtggtgagcgga 420
QY 421 aagtatctattctcgtacgggtgtactccacaggaagacactcccgcaactcccgcc 480
DB 421 aagtatctattctcgtacgggtgtactccacaggaagacactcccgcaactcccgcc 480
QY 481 ctgcgcgatttcaacgggaagtgattcattcgtgctgacgaagacttcgacgca 540
DB 481 ctgcgcgatttcaacgggaagtgattcattcgtgctgacgaagacttcgacgca 540
QY 541 gagggcagagagtcgcctcattcgtgctgacgaagacttcgacgca 600
DB 541 gagggcagagagtcgcctcattcgtgctgacgaagacttcgacgca 600
QY 601 ttggccaaagagctgacgaagtaacattgtttatgcgaagcgaactatgtctgcc 660
DB 601 ttggccaaagagctgacgaagtaacattgtttatgcgaagcgaactatgtctgcc 660
QY 661 atcgccgaacgacgattggaagacgacagacgctggaagcctactacccacg 720
DB 661 atcgccgaacgacgattggaagacgacagacgctggaagcctactacccacg 720
QY 721 ctgtttgaagcagtcgaaagtctcgattggattcccggtccaggaacccgtggttgc 780
DB 721 ctgtttgaagcagtcgaaagtctcgattggattcccggtccaggaacccgtggttgc 780
QY 781 atctttgaagtcagcccgagcagcgagggcctatttcgaagagttgtggagcgtggg 840
DB 781 atctttgaagtcagcccgagcagcgagggcctatttcgaagagttgtggagcgtggg 840
QY 841 gctttcaatttctgtctgcagtcacgagagtcattgtttgacaaagagccacacga 900
DB 841 gctttcaatttctgtctgcagtcacgagagtcattgtttgacaaagagccacacga 900
QY 901 ctggtctagactctggggccaaagacatcattcgtatcgttcattcgttcgaagaga 960
DB 901 ctggtctagactctggggccaaagacatcattcgtatcgttcattcgttcgaagaga 960
```

```
QY 961 gatctcattgggtctctctgagccgctgactgtgttcggtacaaagcgtcccccactggag 1020
DB 961 gatctcattgggtctctctgagccgctgactgtgttcggtacaaagcgtcccccactggag 1020
QY 1021 agcgactactacgaatgctgggaagccgagcgtcgaattgtgaattctagaaacatcg 1080
DB 1021 agcgactactacgaatgctgggaagccgagcgtcgaattgtgaattctagaaacatcg 1080
QY 1081 cccattgtggtcgtgtacaaagacaggtgtgctcttgagtgacggcagcaagagggaatgc 1140
DB 1081 cccattgtggtcgtgtacaaagacaggtgtgctcttgagtgacggcagcaagagggaatgc 1140
QY 1141 gacagatcgtgctggcgacgggtttccgaactttcaactggctc 1184
DB 1141 gacagatcgtgctggcgacgggtttccgaactttcaactggctc 1184
```

## RESULT 12

```
US-09-351-224E-1
; Sequence 1, Application US/09351224E
; GENERAL INFORMATION:
; APPLICANT: Duvick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 5718-111
; CURRENT APPLICATION NUMBER: US/09/351, 224E
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: flavin monooxygenase with intron
US-09-351-224E-1
```

Query Match 72.3%; Score 1184; DB 17; Length 1691;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgtcgccacacagcaactcaagcgattgttcgctcgatcgacgacccatcatcgtt 60
DB 1 atgtcgccacacagcaactcaagcgattgttcgctcgatcgacgacccatcatcgtt 60
QY 61 ggaagcggcctcagcggaatctctgtgtgtacaaatgcgaagactcagactcaacgc 120
DB 61 ggaagcggcctcagcggaatctctgtgtgtacaaatgcgaagactcagactcaacgc 120
QY 121 aaattctcgaggggccccgattttggcggtgtgtgacactgaaacgctacccctggc 180
DB 121 aaattctcgaggggccccgattttggcggtgtgtgacactgaaacgctacccctggc 180
QY 181 gctcgtgtgatttcgagagcgcctctctacaaactgaacattcccgaaagtatgaaagac 240
DB 181 gctcgtgtgatttcgagagcgcctctctacaaactgaacattcccgaaagtatgaaagac 240
QY 241 tggacctgtgttcgctcattcctgacccagaagaggttgcgtcatatgttcacactgt 300
DB 241 tggacctgtgttcgctcattcctgacccagaagaggttgcgtcatatgttcacactgt 300
QY 301 gacagatccggggtctgagaaagacgctctctcggagctgaggtgggtgagcgcg 360
DB 301 gacagatccggggtctgagaaagacgctctctcggagctgaggtgggtgagcgcg 360
QY 361 tatgcagagatctggggccaaagacatcattcgtatcgttcattcgttcgaagaga 420
```



Db 961 gatctatggtctctgagcgcgcgtactggttcggtaccagcgtcccaactggag 1020  
Qy 1021 agcgaactactacgaatgtgtgacaagccgagcgtcgaaattgtgaatctagacaatcg 1080  
Db 1021 agcgaactactacgaatgtgtgacaagccgagcgtcgaaattgtgaatctagacaatcg 1080  
Qy 1081 cccattgtgctattacaagacaggtgtgctcttgatgacgacgacgaagaggaatgc 1140  
Db 1081 cccattgtgctattacaagacaggtgtgctcttgatgacgacgacgaagaggaatgc 1140  
Qy 1141 gacacgatgctgtgagcgcgcgtgttttcgacagtttcaactggctc 1184  
Db 1141 gacacgatgctgtgagcgcgcgtgttttcgacagtttcaactggctc 1184

RESULT 14  
US-09-677-488-1  
; Sequence 1, Application US/09677488  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification  
; FILE REFERENCE: 35718/204100 (5718-111A)  
; CURRENT APPLICATION NUMBER: US/09/677,488  
; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1691  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: flavin monooxygenase with intron  
US-09-677-488-1

Query Match 72.3%; Score 1184; DB 26; Length 1691;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtcggccaccgacactccagagcgatgttccgtcgatcgagcgcatcatcggt 60  
Db 1 atgtcggccaccgacactccagagcgatgttccgtcgatcgagcgcatcatcggt 60  
Qy 61 ggaacggcctcagcgagcgtctgtgtgtacaaattgcgaagctcagactcaacgcc 120  
Db 61 ggaacggcctcagcgagcgtctgtgtgtacaaattgcgaagctcagactcaacgcc 120  
Qy 121 aaatcttcgaggaagcccccgtatttggcggctctgacactggaacgcgtaccctggc 180  
Db 121 aaatcttcgaggaagcccccgtatttggcggctctgacactggaacgcgtaccctggc 180  
Qy 181 gctcgtattgattcgagagcgcctctaccacactgaacattcccggaagtatggaagac 240  
Db 181 gctcgtattgattcgagagcgcctctaccacactgaacattcccggaagtatggaagac 240  
Qy 241 tggacctggcttgcgcgtatctcaccagaagagttgctgtcatatgttccacctgt 300  
Db 241 tggacctggcttgcgcgtatctcaccagaagagttgctgtcatatgttccacctgt 300  
Qy 301 gacaagatccggggcttgagaaagacgtctacttcgagactgaggtgggtgatgcgcgg 360  
Db 301 gacaagatccggggcttgagaaagacgtctacttcgagactgaggtgggtgatgcgcgg 360  
Qy 361 tatccgagatctgggacactggactatcagacgctcggtggtggtgagcgagca 420  
Db 361 tatccgagatctgggacactggactatcagacgctcggtggtggtgagcgagca 420  
Qy 421 aagtatctattctcgtcaggggtgtgtccacggaagacacactccggaactcccgcc 480  
Db 421 aagtatctattctcgtcaggggtgtgtgtccacggaagacacactccggaactcccgcc 480

Qy 481 ctgcgcgattcaacgggaagtgattcattcctgagtcctggcaggaagacttcagcga 540  
Db 481 ctgcgcgattcaacgggaagtgattcattcctgagtcctggcaggaagacttcagcga 540  
Qy 541 gaggggccagagagtgctgcgcgtcatcggtgcggggggccacaaagcatccagattgttcaggag 600  
Db 541 gaggggccagagagtgctgcgcgtcatcggtgcggggggccacaaagcatccagattgttcaggag 600  
Qy 601 ttggccagaagagtgctgcgcgttaaccattattatcagcaagccgagcttatgtctgccc 660  
Db 601 ttggccagaagagtgctgcgcgttaaccattattatcagcaagccgagcttatgtctgccc 660  
Qy 661 atgcggcaacgaacgatgtagaagaaacagacagcctggaaagcctactaccaccacg 720  
Db 661 atgcggcaacgaacgatgtagaagaaacagacagcctggaaagcctactaccaccacg 720  
Qy 721 ctggttgagcagagtcgaaagctctcggattcccggtccagggccacgctcgattggc 780  
Db 721 ctggttgagcagagtcgaaagctctcggattcccggtccagggccacgctcgattggc 780  
Qy 781 atctttgaagtcagcccccagcagcggaggcctatttcgaagagttgtggagcgtggg 840  
Db 781 atctttgaagtcagcccccagcagcggaggcctatttcgaagagttgtggagcgtggg 840  
Qy 841 gcttttaattcttctgcttcgagtcacgagagtcgtgttgacaaaagggccaaaccca 900  
Db 841 gcttttaattcttctgcttcgagtcacgagagtcgtgttgacaaaagggccaaaccca 900  
Qy 901 ctggtctatgactctctggggccaaaagactcgtatcgtatcgtcgaatccgcaagaga 960  
Db 901 ctggtctatgactctctggggccaaaagactcgtatcgtatcgtcgaatccgcaagaga 960  
Qy 961 gatctcatggtctcttgagcgcctactggttcgggtacccaagcgtcccccactggag 1020  
Db 961 gatctcatggtctcttgagcgcctactggttcgggtacccaagcgtcccccactggag 1020  
Qy 1021 agcgaactactcgaatgtctggacagccgagcgtcgaaattgtgaatctagaacaatcg 1080  
Db 1021 agcgaactactcgaatgtctggacagccgagcgtcgaaattgtgaatctagaacaatcg 1080  
Qy 1081 cccattgtgctgtttacaaagacagcagggtgtgctcttctgagtcagcggcagcaggggaatgc 1140  
Db 1081 cccattgtgctgtttacaaagacagcagggtgtgctcttctgagtcagcggcagcaggggaatgc 1140  
Qy 1141 gacacgatcgtgctggcagcgggttttcgacagtttcaactggctc 1184  
Db 1141 gacacgatcgtgctggcagcgggttttcgacagtttcaactggctc 1184

RESULT 15  
US-09-677-488A-1  
; Sequence 1, Application US/09677488A  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 35718/204100  
; CURRENT APPLICATION NUMBER: US/09/677,488A  
; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1691  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc\_feature



```

; LOCATION: (0)...(0)
; OTHER INFORMATION: flavin monooxygenase with intron
US-09-677-488A-1

```

Query Match	72.3%	Score 1184	DB 26	Length 1691
Best Local Similarity	100.0%	Prod. No. 0		
Matches 1184	Conservative	0	Mismatches	0
			Indels	Gaps
QY 1	atgtcggccacagcaactccagagcgcgattgttccgtgcgcatgcgacgcatcatcgtt	60		
DB 1	atgtcggccacagcaactccagagcgcgattgttccgtgcgcatgcgacgcatcatcgtt	60		
QY 61	ggagcgcgctccagcgcatctctgtgtgacaaattgcgaaagctcagactcaacgc	120		
DB 61	ggagcgcgctccagcgcatctctgtgtgacaaattgcgaaagctcagactcaacgc	120		
QY 121	aaaattctcaagagagcccccatttttgcgcgctctgacactgaaacgctacccctgc	180		
DB 121	aaaattctcaagagagcccccatttttgcgcgctctgacactgaaacgctacccctgc	180		
QY 181	gtctgtgtgattcggagcgcctcttaccacatccgaattcccgaaagtatgaaagac	240		
DB 181	gtctgtgtgattcggagcgcctcttaccacatccgaattcccgaaagtatgaaagac	240		
QY 241	tggacctgttcttcgcttatctcaccagaaaagtgtgttcataattgccacctgt	300		
DB 241	tggacctgttcttcgcttatctcaccagaaaagtgtgttcataattgccacctgt	300		
QY 301	gacaaatccggggcttgagaaaagccttacttcggagctgagtggtgtgatgcgcg	360		
DB 301	gacaaatccggggcttgagaaaagccttacttcggagctgagtggtgtgatgcgcg	360		
QY 361	tatgcagagatcttgggcaactcggactgtcaagacgttcgcttgcccatgttcgacggca	420		
DB 361	tatgcagagatcttgggcaactcggactgtcaagacgttcgcttgcccatgttcgacggca	420		
QY 421	aagtatctcatcttcgctacgggtgtgtccacagggagcacactccgcgactcccgcg	480		
DB 421	aagtatctcatcttcgctacgggtgtgtccacagggagcacactccgcgactcccgcg	480		
QY 481	ctgcgcgatttcaacgggaaggttgattcatctcgagtccttggcagcaagaacttcgacgca	540		
DB 481	ctgcgcgatttcaacgggaaggttgattcatctcgagtccttggcagcaagaacttcgacgca	540		
QY 541	gaaggccagagatcgcgcgtcatcagttgccgggcccacagcatcagatgttccagag	600		
DB 541	gaaggccagagatcgcgcgtcatcagttgccgggcccacagcatcagatgttccagag	600		
QY 601	tggccaaagaggtgcgcgtcatcagttgccgggcccacagcatcagatgttccagag	660		
DB 601	tggccaaagaggtgcgcgtcatcagttgccgggcccacagcatcagatgttccagag	660		
QY 661	atcgggcaacgaaacagatggataggaacgaacagacagccttggaaagccttaccaccaag	720		
DB 661	atcgggcaacgaaacagatggataggaacgaacagacagccttggaaagccttaccaccaag	720		
QY 721	ctgtttgaagcgagtcgaaagtctcgatttgagttcccggtccagagcagcgtcgattggc	780		
DB 721	ctgtttgaagcgagtcgaaagtctcgatttgagttcccggtccagagcagcgtcgattggc	780		
QY 781	atctttgaagtcagcccccgagcagcgagggcctatttcgaagaagttgtggagagcgtgg	840		
DB 781	atctttgaagtcagcccccgagcagcgagggcctatttcgaagaagttgtggagagcgtgg	840		
QY 841	gcctttaattttctgtgtccagtagtcacgagaagtcattgtgacaaaagcccaaccca	900		
DB 841	gcctttaattttctgtgtccagtagtcacgagaagtcattgtgacaaaagcccaaccca	900		
QY 901	ctggtctatgacctctgggcacaaagacatcgatctcgtatcgtcaactccggcaagaga	960		
DB 901	ctggtctatgacctctgggcacaaagacatcgatctcgtatcgtcaactccggcaagaga	960		



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 01:38:33 ; Search time 151.01 Seconds  
(without alignments)  
2664.378 Million cell updates/sec

Title: US-09-882-694A-2  
Perfect score: 1638  
Sequence: 1 atgtcggccaccagcaactc.....tgcttcattgctccctgctaa 1638

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	1.2	1814	1	US-08-579-667-7
C 2	17	1.0	453	1	US-08-562-985A-8
C 3	17	1.0	845	1	US-08-266-451B-1
C 4	17	1.0	845	2	US-08-748-725-1
C 5	17	1.0	5889	1	US-08-463-092B-5
C 6	17	1.0	5889	2	US-08-462-109A-5
C 7	17	1.0	5889	2	US-08-460-907B-5
C 8	17	1.0	5889	3	US-08-463-179A-5
C 9	17	1.0	5889	3	US-08-461-384B-5
C 10	17	1.0	8224	3	5180808-1
C 11	17	1.0	35060	3	US-08-814-095-7
C 12	17	1.0	90050	4	US-09-245-041-5
C 13	17	1.0	4411529	4	US-09-103-840A-1
C 14	16	1.0	110	1	US-08-465-798-2
C 15	16	1.0	115	1	US-08-465-798-1
C 16	16	1.0	461	1	US-08-133-179-2
C 17	16	1.0	461	1	US-08-133-011-117
C 18	16	1.0	461	1	US-08-322-730A-117
C 19	16	1.0	461	1	US-08-387-874-90
C 20	16	1.0	461	2	US-08-383-618-117
C 21	16	1.0	461	4	US-08-907-739-117
C 22	16	1.0	461	5	PCT-US93-08364-90
C 23	16	1.0	470	4	US-09-025-769B-286
C 24	16	1.0	472	4	US-08-820-931-5
C 25	16	1.0	476	3	US-08-050-259B-19
C 26	16	1.0	502	4	US-08-998-416-124
C 27	16	1.0	507	4	US-09-489-292-1

28	16	1.0	686	3	US-08-370-223-6	Sequence 6, Appli
C 29	16	1.0	698	4	US-08-998-416-54	Sequence 54, Appl
C 30	16	1.0	713	4	US-08-998-416-135	Sequence 135, App
C 31	16	1.0	798	4	US-09-007-119-17	Sequence 17, Appl
C 32	16	1.0	801	4	US-09-007-119-16	Sequence 16, Appl
C 33	16	1.0	930	1	US-08-096-182A-1	Sequence 1, Appli
C 34	16	1.0	930	1	US-08-877-109-1	Sequence 1, Appli
C 35	16	1.0	930	3	US-08-798-760-1	Sequence 1, Appli
C 36	16	1.0	930	5	PCT-US94-08327-1	Sequence 1, Appli
C 37	16	1.0	1015	2	US-08-820-178A-6	Sequence 6, Appli
C 38	16	1.0	1015	3	US-09-055-699-6	Sequence 6, Appli
C 39	16	1.0	1015	4	US-09-273-565-6	Sequence 6, Appli
C 40	16	1.0	1015	4	US-09-565-538-6	Sequence 6, Appli
C 41	16	1.0	1200	4	US-08-818-112-11	Sequence 11, Appl
C 42	16	1.0	1200	4	US-08-818-111-11	Sequence 11, Appl
C 43	16	1.0	1200	4	US-09-056-556-11	Sequence 11, Appl
C 44	16	1.0	1217	1	US-08-726-462B-1	Sequence 1, Appli
C 45	16	1.0	1217	2	US-08-642-330-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-579-667-7/c  
; Sequence 7, Application US/08579667  
; Patent No. 5705624  
; GENERAL INFORMATION:  
; APPLICANT: Pitmaurice, Wayne P.  
; APPLICANT: Hellmann, Gary M.  
; APPLICANT: Grill, Laurence K.  
; APPLICANT: Kumagai, Monto H.  
; APPLICANT: Della-Cioppa, Guy R.  
; TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN  
; TITLE OF INVENTION: PHYTOENE BIOSYNTHESIS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Virginia C. Bennett  
; STREET: 1211 East Morehead Street, PO Drawer 34009  
; CITY: Charlotte  
; STATE: No. 5705624th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/579,667  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Virginia C.  
; REGISTRATION NUMBER: 37,092  
; REFERENCE/SOCKET NUMBER: 827-196  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-420-2200  
; TELEFAX: 919-881-3175  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1814 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 363..1592  
US-08-579-667-7

Query Match 1.2% Score 19; DB 1; Length 1814;

```
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 843 cttaattttctgtctgc 861
Db 66 CTTAATTTCTGTCTGC 48

RESULT 2
US-08-562-985A-8/c
; Sequence 8, Application US/08562985A
; Patent No. 5821057
; GENERAL INFORMATION:
; APPLICANT: CHIANG, John Y.L.
; TITLE OF INVENTION: AN ASSAY FOR AGENTS THAT AFFECT
; TITLE OF INVENTION: CHOLESTEROL 7ALPHA-HYDROXYLASE EXPRESSION AND A
; TITLE OF INVENTION: CHARACTERIZATION OF ITS REGULATORY ELEMENTS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/562.985A
; FILING DATE: 27-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 18748/213/HOCE
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 8:
; LENGTH: 453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-562-985A-8

Query Match 1.0%; Score 17; DB 1; Length 453;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1571 atgacgagatctgtcaa 1587
Db 128 ATGACGAGATCTGTCAA 112

RESULT 3
US-08-266-451B-1/c
; Sequence 1, Application US/08266451B
; Patent No. 5623054
; GENERAL INFORMATION:
; APPLICANT: Zhang et al.
; TITLE OF INVENTION: CRUCIFER APT PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION NUMBER: US/08/266.451B
; FILING DATE: 23-June-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/266.451
; FILING DATE: 23-June-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/219002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070

Best Local Similarity 100.0%; Pred. No. 44;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 gaaagctcagactcaac 117
Db 333 GAAAGCTCAGACTCAAC 317

RESULT 4
US-08-748-725-1/c
; Sequence 1, Application US/08748725
; Patent No. 5859346
; GENERAL INFORMATION:
; APPLICANT: Zhang et al.
; TITLE OF INVENTION: CRUCIFER APT PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION NUMBER: US/08/748.725
; FILING DATE: 23-June-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/266.451
; FILING DATE: 23-June-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/219002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
```

TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 845  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-748-725-1

Query Match 1.0%; Score 17; DB 2; Length 845;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 gaagctcagactcaac 117  
|||||  
DB 333 GAAGCTCAGACTCAAC 317

RESULT 5  
US-08-463-092B-5  
; Sequence 5, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; STREET: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: 01546  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5889 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6 4589  
US-08-463-092B-5

Query Match 1.0%; Score 17; DB 1; Length 5889;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 gaagctcactacccca 718  
|||||  
DB 3554 GAAGGCTACTACCCCA 3570

RESULT 6  
US-08-462-109A-5  
; Sequence 5, Application US/08462109A  
; Patent No. 5882875  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,109A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: PQI-002CP4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5889 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 6 4589  
US-08-462-109A-5



Patent No. 6025473  
GENERAL INFORMATION:  
APPLICANT: Cole, Susan P.C.  
APPLICANT: Deeley, Roger G.  
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS  
STREET: Queen's University at Kingston  
CITY: Kingston  
STATE: Ontario  
COUNTRY: CANADA  
ZIP: K7L 3N6  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,384B  
FILING DATE: 05-JUN-95  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: 07/966,923  
FILING DATE: 27-OCT-1992  
APPLICATION NUMBER: 08/029,340  
FILING DATE: 8-MAR-1993  
APPLICATION NUMBER: 08/141,893  
FILING DATE: 26-OCT-1993  
APPLICATION NUMBER: 08/407,207  
FILING DATE: 20-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Steeg, Carol Miernicki  
REGISTRATION NUMBER: 39,539  
REFERENCE/DOCKET NUMBER: Q1547  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 545-2342  
TELEFAX: (613) 545-6853  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5889 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6 4589  
US-08-461-384B-5

Query Match 1.0%; Score 17; DB 3; Length 5889;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 gaaggcctactaccoca 718  
|||||

DB 3554 GAAGGCGCTACTACCCCA 3570

RESULT 10  
5180808-1  
Patent No. 5180808  
APPLICANT: RUOSLAHTI, ERKKI I.  
TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID  
SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN  
ANTIBODIES, AND METHODS OF DETECTING THE SAME  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/441,179  
FILING DATE: 27-NOV-1989  
SEQ ID NO: 1:  
LENGTH: 8224  
5180808-1

Query Match 1.0%; Score 17; DB 6; Length 8224;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 atcatcggttgagccgg 68  
|||||

DB 7456 atcatcggttgagccgg 7472

RESULT 11  
US-08-814-095-7  
Sequence 7, Application US/08814095  
Patent No. 6025183  
GENERAL INFORMATION:  
APPLICANT: Soreq, Hermona  
APPLICANT: Zakut, Haim  
APPLICANT: Shaul, Moshe  
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR  
ANTI-CHOLINESTERASE SUBSTANCES  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KOHN & ASSOCIATES  
STREET: 30500 No. 6025183thwestern Highway, Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/814,095  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Montgomery, Ilene N.  
REGISTRATION NUMBER: 38,972  
REFERENCE/DOCKET NUMBER: 2391.00066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (248) 539-5050  
TELEFAX: (248) 539-5055  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc= "Cosmid including ACHE"  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 7q22  
FEATURE:  
NAME/KEY: promoter  
LOCATION: 4089..22464  
OTHER INFORMATION: /function= "ACHE Promotor"  
OTHER INFORMATION: /standard\_name= "ACHE Promotor"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 22465..22537  
OTHER INFORMATION: /function= "non-translated"  
OTHER INFORMATION: /gene= "ACHE"  
OTHER INFORMATION: /number= 1  
FEATURE:  
NAME/KEY: exon

```
; LOCATION: 24090..25177 experimental
; IDENTIFICATION METHOD: /function= "(translation start:
; OTHER INFORMATION: /function= "24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 27385..27387
; FEATURE:
; NAME/KEY: exon
; LOCATION: 28008..28129
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 28129..28131
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34528..34895)
; OTHER INFORMATION: /function= "arsenite resistance
; OTHER INFORMATION: gene"
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34092..34358)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33779..33963)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33493..33591)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33297..33408)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32959..33094)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 6
```

```
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32569..32628)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32386..32468)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31894..32080)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31363..31534)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31131..31284)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
; US-08-814-095--7
```

Query Match 1.0%; Score 17; DB 3; Length 35060;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 150 cgcgcctctggcactgga 166  
|||||  
Db 22264 CGCGCTCTGGCCTGGA 22280

## RESULT 12

US-09-245-041-5/c  
; Sequence 5, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041



```
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,530
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

Query Match          1.0%; Score 17; DB 4; Length 90050;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 cctctaccactgaac 219
|||||
DB 348 CCTCTACCACTGAAC 332

RESULT 13
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match          1.0%; Score 17; DB 4; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 483 cgcgcatttcaacggga 499
|||||
DB 3223359 CGCGGATTTCACGGGA 3223343

RESULT 14
US-08-465-798-2
; Sequence 2, Application US/08465798
; Patent No. 5683877
; GENERAL INFORMATION:
; APPLICANT: LU-CHANG, A-Lein
; APPLICANT: HSU, Jh-Chang
; TITLE OF INVENTION: DETECTION OF GENE MUTATIONS
; TITLE OF INVENTION: WITH MISMATCH REPAIR ENZYMES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP 20037
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/465,798
; APPLICATION NUMBER: US/08/465,798
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/309,629
; FILING DATE: 21-SEPT-1994
; PRIOR APPLICATION DATA:
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KIT, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6013-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: synthetic DNA
; HYPOTHETICAL: NO
US-08-465-798-2

Query Match          1.0%; Score 16; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 agggagcccccgattt 146
|||||
DB 34 AGGGAGCCCCGATTT 49

RESULT 15
US-08-465-798-1/c
; Sequence 1, Application US/08465798
; Patent No. 5683877
; GENERAL INFORMATION:
; APPLICANT: LU-CHANG, A-Lein
; APPLICANT: HSU, Jh-Chang
; TITLE OF INVENTION: DETECTION OF GENE MUTATIONS
; TITLE OF INVENTION: WITH MISMATCH REPAIR ENZYMES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/465,798
; APPLICATION NUMBER: US/08/465,798
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/309,629
; FILING DATE: 21-SEPT-1994
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 07/859,072
```

```

; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KIT, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6013-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: synthetic DNA
; HYPOTHETICAL: NO
; US-08-465-798-1

```

```

Query Match      1.0%  Score 16; DB 1; Length 115;
Best Local Similarity 100.0%  Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 131 agggagccccgattt 146
Db 86 AGGGAGCCCCCGATT 71

```

Search completed: April 27, 2002, 05:54:18  
Job time: 15345 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 01:38:37 ; Search time 705 Seconds  
(without alignments)  
3989.086 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638  
Sequence: 1 atgtggccaccagcaactc.....tgcttcatactccctgctaa 1638

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_032802:\*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1638	100.0	1638	22	Exophiala spinifer
2	1184	72.3	1691	22	Exophiala spinifer
3	19	1.2	1814	19	Phytoene synthase
4	19	1.2	21100	22	Human nervous syst
5	19	1.2	349980	22	Pyrococcus abyssi
6	18	1.1	432	21	Xenopus secreted e
7	18	1.1	1100	21	Human secreted pro
8	18	1.1	1110	16	Type I matrix meta
9	18	1.1	1158	23	Drosophila melanog

10	18	1.1	1494	22	AA889186	Human oxidoreducta
11	18	1.1	1549	22	AA894526	Human full-length
12	18	1.1	2095	23	ABL15755	Drosophila melanog
13	18	1.1	2171	21	AA776642	Human ORFX ORF2197
14	18	1.1	2624	22	AAH17299	Human CDNA sequenc
15	18	1.1	3158	23	ABL20376	Drosophila melanog
16	18	1.1	3414	23	ABL10835	Drosophila melanog
17	18	1.1	4160	23	ABL15754	Drosophila melanog
18	18	1.1	4857	23	ABL03544	Drosophila melanog
19	18	1.1	6656	23	ABL13932	Drosophila melanog
20	18	1.1	12289	23	ABL10834	Drosophila melanog
21	18	1.1	349980	22	AAH41224	Pyrococcus abyssi
22	18	1.0	319	22	AA538063	Novel human diapo
23	17	1.0	362	21	AA97754	Plantomyces maris
24	17	1.0	381	21	AA612316	Human secreted pro
25	17	1.0	396	23	ABL24309	Drosophila melanog
26	17	1.0	407	21	AA37235	Arabidopsis thalia
27	17	1.0	456	21	AAA56852	Mycoplasma hyopneu
28	17	1.0	463	22	ABA42792	Human breast cell
29	17	1.0	463	22	ABA53220	Human foetal liver
30	17	1.0	463	22	ABA22995	Human brain expres
31	17	1.0	463	22	AAK01468	Human bone marrow
32	17	1.0	463	22	AAK26916	Human bone marrow
33	17	1.0	463	22	AA111538	Probe #1471 for ge
34	17	1.0	463	22	AA132820	Probe #1506 used t
35	17	1.0	463	22	AA101456	Probe #1447 used t
36	17	1.0	477	22	AA52889	Human foetal liver
37	17	1.0	477	22	ABA22671	Probe #1137 for ge
38	17	1.0	477	22	AAK26600	Human bone marrow
39	17	1.0	477	22	AA132492	Probe #1178 used t
40	17	1.0	481	21	AA000422	Human MINT4 DNA ci
41	17	1.0	576	21	AAK00304	Arabidopsis SERK 1
42	17	1.0	735	22	AAK92137	Human CDNA 5'-end
43	17	1.0	735	22	AAK93398	Human CDNA clone r
44	17	1.0	845	17	AA111323	DNA encoding recom
45	17	1.0	917	23	AA577521	DNA encoding novel

#### ALIGNMENTS

RESULT 1  
AAD02691  
ID AAD02691 standard; DNA; 1638 BP.  
XX AAD02691:  
XX AC  
XX 02-MAY-2001 (first entry)  
XX Exophiala spinifera flavin monooxygenase coding sequence.  
XX Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
KW aldehyde dehydrogenase; perasease; p-glycoprotein; fumonisin esterase;  
KW detoxification; mycotoxin; animal feed; human feed; silage;  
KW transgenic plant; transgenic animal; microbial spray; ds.  
XX Exophiala spinifera.  
OS  
XX Key Location/Qualifiers  
XX CDS 1..1638  
XX FT /\*tag= a  
XX FT /product= "E. spinifera flavin monooxygenase"  
XX WO200105980-A1.  
XX 25-JAN-2001.  
XX PF 14-JUL-1999; 99WO-US15824.  
XX XX  
XX PR 14-JUL-1999; 99WO-US15824.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX (CURA-) CURAGEN CORP.

XX The patent discloses novel polynucleotides encoding *Exophiala spinifera*  
PI funiculin degradative or transport enzymes such as flavin monooxygenase,  
XX aldehyde dehydrogenase, peroxidase and p-glycoprotein. A funiculin esterase  
XX enzyme with at least one funiculin degradative enzyme is useful for  
XX detoxifying funiculin or a structurally related mycotoxin present in  
XX harvested grain, on application to a plant or to harvested grain  
XX during processing, or to processed grain that is to be used as animal  
XX or human feed, or as a silage. Nucleotide fragments of the present  
XX invention are useful as probes and primers. They can be introduced  
XX into microorganisms that multiply on plants to deliver enzymes to  
XX potential target crops. The genes encoding the degradative enzymes are  
XX introduced via a vector into a microbial host and the transformed host  
XX is supplied to the environment, plants or animals for reducing the  
XX pathogenicity of a fungus producing funiculin. The genes of the  
XX invention are fermented in a bacterial host and the resulting bacteria  
XX is processed and used as a microbial spray. The nucleotide sequences  
XX can be used alone or in combination to engineer microbes or other  
XX organisms to metabolize funiculin and resist its toxic effects.  
XX The present DNA sequence is the flavin monooxygenase coding sequence  
XX from *Exophiala spinifera*.  
XX  
XX Sequence 1638 BP; 388 A; 432 C; 462 G; 356 T; 0 other;

QY 481 ctccgacgatttcaacggggaaggtgatttcatttcagtgctgacacgaagacttcagcga 540  
DB 481 ctccgacgatttcaacggggaaggtgatttcatttcagtgctgacacgaagacttcagcga 540  
QY 541 gagggccagagagtcgcgcgtcaatcggtgcggggccacaaagcatccagatgttccagag 600  
DB 541 gagggccagagagtcgcgcgtcaatcggtgcggggccacaaagcatccagatgttccagag 600  
QY 601 ttggccaaagaggtgcacccaggttaacattgttattgcgaagcccgatgtctgtgcc 660  
DB 601 ttggccaaagaggtgcacccaggttaacattgttattgcgaagcccgatgtctgtgcc 660  
QY 661 atgcccgaacgaatgtagtagaagcaacagacagcctggaagccctactaccacacg 720  
DB 661 atgcccgaacgaatgtagtagaagcaacagacagcctggaagccctactaccacacg 720  
QY 721 ctgtttgaagcagatcgaaggtctcggtattggatttcctccggtccagggcctcggtggc 780  
DB 721 ctgtttgaagcagatcgaaggtctcggtattggatttcctccggtccagggcctcggtggc 780  
QY 781 atctttgaagtcagccccgagcagcgagggtcctatttcgaagagcttggtggagcgtggg 840  
DB 781 atctttgaagtcagccccgagcagcgagggtcctatttcgaagagcttggtggagcgtggg 840  
QY 841 gcttttaattttctgtgtccagtcacccagagagtcagtcgttgcacaaagggccacacga 900  
DB 841 gcttttaattttctgtgtccagtcacccagagagtcagtcgttgcacaaagggccacacga 900  
QY 901 ctgggtctatgactctctggccaaagagtcgactctgctatcgctcaatcccggaagaga 960  
DB 901 ctgggtctatgactctctggccaaagagtcgactctgctatcgctcaatcccggaagaga 960  
QY 961 gattctcagtcctctggaagcccgctactggttcggtacacacgctccacactggag 1020  
DB 961 gattctcagtcctctggaagcccgctactggttcggtacacacgctccacactggag 1020  
QY 1021 agcactactacgaatgctggacaagccgagcgctcgaaattgtgaatctagaacacacg 1080  
DB 1021 agcactactacgaatgctggacaagccgagcgctcgaaattgtgaatctagaacacacg 1080  
QY 1081 cccattgtggctgtttacaagagacaggtgtgctcttgagtgacggcgagcaagggaaatgc 1140  
DB 1081 cccattgtggctgtttacaagagacaggtgtgctcttgagtgacggcgagcaagggaaatgc 1140  
QY 1141 gacacgactcgtgctggcgagcggttttcgacagtttcactggctcattgacacatatggc 1200  
DB 1141 gacacgactcgtgctggcgagcggttttcgacagtttcactggctcattgacacatatggc 1200  
QY 1201 ttgaaaaaacaagcagcgagtggaacctgaagggagtgtagaagatggcatatctactat 1260  
DB 1201 ttgaaaaaacaagcagcgagtggaacctgaagggagtgtagaagatggcatatctactat 1260  
QY 1261 atggagtgcttctctatggcttcccaatgcttcttctgctgcacagcgctcaagcccg 1320  
DB 1261 atggagtgcttctctatggcttcccaatgcttcttctgctgcacagcgctcaagcccg 1320  
QY 1321 accgtcccttcccaagcgcccaacgacatagaaccccaagtcgacttgatcgcgatata 1380  
DB 1321 accgtcccttcccaagcgcccaacgacatagaaccccaagtcgacttgatcgcgatata 1380  
QY 1381 attgcaagttggagcgccgagcgcacgtccgtttgagggcagcaaatcagcaacagag 1440  
DB 1381 attgcaagttggagcgccgagcgcacgtccgtttgagggcagcaaatcagcaacagag 1440  
QY 1441 gcatggtcgattatgattgcgaagatgaacagacacactcttcccttgcagggattcg 1500  
DB 1441 gcatggtcgattatgattgcgaagatgaacagacacactcttcccttgcagggattcg 1500  
QY 1501 tgggtgactggaggaacacatccctgggaagcaacacgtgtctttacactcataggcg 1560  
DB 1501 tgggtgactggaggaacacatccctgggaagcaacacgtgtctttacactcataggcg 1560  
QY 1561 attgctctctatgagcagatctgtcaagagaggttgcccaattgggattgtatgtg 1620

Db 1561 attgctctctatgagcagatctgtcaagaaggtggccaattggatgggtttgatgtg 1620

Qy 1621 ettatgctctctgctctaa 1638

Db 1621 ettatgctctctgctctaa 1638

RESULT 2

AD02690 ID AAD02690 standard; DNA: 1691 BP.

XX AAD02690;

XX 02-MAY-2001 (first entry)

XX Exophiala spinifera flavin monooxygenase genomic DNA.

KW Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
KW aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;  
KW detoxification; mycotoxin; animal feed; human feed; silage;  
KW transgenic plant; transgenic animal; microbial spray; ds.

XX Exophiala spinifera.

OS WO200105980-A1.

XX 25-JAN-2001.

XX 14-JUL-1999; 99WO-US15824.

XX 14-JUL-1999; 99WO-US15824.

XX (PION-) PIONEER HI-BRED INT INC.

XX (CURA-) CURAGEN CORP.

PI Duwick JP, Maddox J, Gilliam J, Folkerts O, Crastra OR;

XX WPI: 2001-147345/15.

DR P-PSDB; AAY72634.

XX Novel polynucleotides encoding Exophiala degradative or transport  
PT enzyme which is useful for detoxifying fumonisin or structurally  
PT related mycotoxin during processing of grain for human or animal food  
PT consumption -

PS Claim 1b; page 63-64; 90pp; English.

XX The patent discloses novel polynucleotides encoding Exophiala spinifera  
CC fumonisin degradative or transport enzymes such as flavin monooxygenase,  
CC aldehyde dehydrogenase, permease and p-glycoprotein. A fumonisin esterase  
CC enzyme with at least one fumonisin degradative enzyme is useful for  
CC detoxifying fumonisin or a structurally related mycotoxin present in  
CC harvested grain, on application to a plant or to harvested grain  
CC during processing, or to processed grain that is to be used as animal  
CC or human feed, or as a silage. Nucleotide fragments of the present  
CC invention are useful as probes and primers. They can be introduced  
CC into microorganisms that multiply on plants to deliver enzymes to  
CC potential target crops. The genes encoding the degrading enzymes are  
CC introduced via a vector into a microbial host and the transformed host  
CC is supplied to the environment, plants or animals for reducing the  
CC pathogenicity of a fungus producing fumonisin. The genes of the  
CC invention are fermented in a bacterial host and the resulting bacteria  
CC is processed and used as a microbial spray. The nucleotide sequences  
CC can be used alone or in combination to engineer microbes or other  
CC organisms to metabolise fumonisin and resist its toxic effects.  
CC The present sequence is a genomic DNA encoding Exophiala spinifera  
CC flavin monooxygenase, a fumonisin degradative enzyme.

XX Sequence 1691 BP; 397 A; 444 C; 480 G; 370 T; 0 other;

Query Match

72.3%; Score 1184; DB 22; Length 1691;

		Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps	0;
		Matches 1184;	Conservative	0;						
Qy	1	atgtcgccacagcaaacctccagaggcgattgttcggtcgatcgacgcacatcatt	60							
Db	1	atgtcgccacagcaaacctccagaggcgattgttcggtcgatcgacgcacatcatt	60							
Qy	61	ggagccggcctcagcggcatctctgtgtacaaaatgcaaaagctcagactcaacgcc	120							
Db	61	ggagccggcctcagcggcatctctgtgtacaaaatgcaaaagctcagactcaacgcc	120							
Qy	121	aaatcttcgaggagccccgatttggcggtctggcaactggaaacgctaccctggc	180							
Db	121	aaatcttcgaggagccccgatttggcggtctggcaactggaaacgctaccctggc	180							
Qy	181	gtctgttgattcgagagcccttccaaactgaactcccaagattccgaagatggaaagac	240							
Db	181	gtctgttgattcgagagcccttccaaactgaactcccaagattccgaagatggaaagac	240							
Qy	241	tgaccctgtcttcgctatctcctgacacagaaagattgtctcatatttccaccactgt	300							
Db	241	tgaccctgtcttcgctatctcctgacacagaaagattgtctcatatttccaccactgt	300							
Qy	301	gacaagatccgggcttgagaaaagactctacttcgagctgagtggtgtgatgcgag	360							
Db	301	gacaagatccgggcttgagaaaagactctacttcgagctgagtggtgtgatgcgag	360							
Qy	361	tatgcagagatctgggcaactgactgtcaagacgtcggtgagccatgttgcagcgca	420							
Db	361	tatgcagagatctgggcaactgactgtcaagacgtcggtgagccatgttgcagcgca	420							
Qy	421	aagtattctctctgctacgggtgtctccacaggaagcacactcccgcaactcccgcc	480							
Db	421	aagtattctctctgctacgggtgtctccacaggaagcacactcccgcaactcccgcc	480							
Qy	481	ctcgcgatttcaacgggaagtgatttcctcagtgcttgcctggcagaaacttccagca	540							
Db	481	ctcgcgatttcaacgggaagtgatttcctcagtgcttgcctggcagaaacttccagca	540							
Qy	541	gagggccagagatcgccgctcattcggttcggggcccaagaagcagatttccagag	600							
Db	541	gagggccagagatcgccgctcattcggttcggggcccaagaagcagatttccagag	600							
Qy	601	ttggccaaagaagctgacaggaataccatgtttatgcgaagggcgagctattgtgcc	660							
Db	601	ttggccaaagaagctgacaggaataccatgtttatgcgaagggcgagctattgtgcc	660							
Qy	661	atgcggcaagacatgtagaagcaacagacagcctggagcctactaccaccag	720							
Db	661	atgcggcaagacatgtagaagcaacagacagcctggagcctactaccaccag	720							
Qy	721	ctgtttgaagcagtcgaaagtctcggttggattcccggtccagcgacgctggttgac	780							
Db	721	ctgtttgaagcagtcgaaagtctcggttggattcccggtccagcgacgctggttgac	780							
Qy	781	atctttgaagtcagcccgagcagcgaggagcctatttcgaagagttgtggagcgtgg	840							
Db	781	atctttgaagtcagcccgagcagcgaggagcctatttcgaagagttgtggagcgtgg	840							
Qy	841	gcctttaatttctgttcgagtcacagagaagtcattgtacaaaaggcccaaccga	900							
Db	841	gcctttaatttctgttcgagtcacagagaagtcattgtacaaaaggcccaaccga	900							
Qy	901	ctgggtctatgactctggggccaaaagactcgatctcgatcgatccatccgcaagaga	960							
Db	901	ctgggtctatgactctggggccaaaagactcgatctcgatcgatccatccgcaagaga	960							
Qy	961	gatctatggtctcttgagcgcctactggttcggtaccagcgtcccccactggag	1020							
Db	961	gatctatggtctcttgagcgcctactggttcggtaccagcgtcccccactggag	1020							
Qy	1021	agcgatactatgagaatctggacagccgctcgaaattgtgaatcgaacaatcg	1080							
Db	1021	agcgatactatgagaatctggacagccgctcgaaattgtgaatcgaacaatcg	1080							

Db 1021 agcgactactaagaaatgctgacaaagcgcgagctgcgaaattgtgaattctagaacaatcg 1080  
Qy 1081 cccattgtgctgttacaagaacaggtgtgctcttgaagtacgacgacgaggggaatgc 1140  
Db 1081 cccattgtgctgttacaagaacaggtgtgctcttgaagtacgacgacgaggggaatgc 1140  
Qy 1141 gacacgatgctgtgacgacgaggtttcgcacagtttccactgctc 1184  
Db 1141 gacacgatgctgtgacgacgaggtttcgcacagtttccactgctc 1184

RESULT 3  
AAV03881/c  
ID AAV03881 standard; cDNA; 1814 BP.  
XX AAV03881;  
XX 29-APR-1998 (first entry)  
XX Phytoene synthase coding sequence from N. tabacum.  
XX Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;  
XX ultra violet absorber; food colour; ss.  
XX Nicotiana tabacum.

XX Key Location/Qualifiers  
XX CDS 363..1595  
XX /\*tag= a  
XX US5705624-A.  
XX 06-JAN-1998.  
XX 27-DEC-1995; 95US-0579667.  
XX 27-DEC-1995; 95US-0579667.  
XX (DELLA) DELLA-CIOPPA G R.  
XX (FITZ) FITZMAURICE W P.  
XX (GRILL) GRILL L K.  
XX (HELL) HELLMANN G M.  
XX (KUMA) KUNAGAI M H.

XX Della-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GM;  
XX Kumagai MH;  
XX WPI; 1998-086196/08.  
XX P-PSDB; AAW41060.  
XX DNA encoding tobacco phytoene synthase polypeptides - useful for  
XX producing recombinant polypeptides or transgenic plants  
XX Claim 1; Column 33-36; 25pp; English.

XX This sequence encodes the phytoene synthetase from Nicotiana tabacum.  
XX The phytoene synthetase coding sequence represents a cDNA of the  
XX invention. The isolated nucleic acid molecules are used for producing  
XX recombinant polypeptides or transgenic plants with enhanced ability to  
XX synthesise carotenoids. Phytoene has been used as a ultra violet absorber  
XX and other carotenoids have been used as food colours, animal feeds and in  
XX the pharmaceutical and cosmetics industries.  
XX Sequence 1814 BP; 577 A; 281 C; 441 G; 515 T; 0 other;

Query Match 1.2%; Score 19; DB 19; Length 1814;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 843 ctttaattttctgttgc 861  
Db 66 CTTTAATTTCTGCTGC 48

RESULT 4  
ABA15497/c  
ID ABA15497 standard; DNA; 21100 BP.  
XX ABA15497;  
XX 23-JAN-2002 (first entry)  
XX Human nervous system related polynucleotide SEQ ID NO 7828.  
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
XX antiparkinsonian; antislaking; antianemic; antiarthritic; cancer;  
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
XX antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;  
XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX Homo sapiens.  
XX WO200159063-A2.  
XX 16-AUG-2001.  
XX 17-JAN-2001; 2001WO-US01334.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 14-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226688.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234224.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234597.  
 PR 25-SEP-2000; 2000US-0234598.  
 PR 26-SEP-2000; 2000US-0234584.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235835.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239335.  
 PR 13-OCT-2000; 2000US-0239337.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 20-OCT-2000; 2000US-0242221.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 01-DEC-2000; 2000US-0251160.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI  
 XX WPI; 2001-541565/60.  
 DR  
 XX  
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system  
 PT cancers and metastases -  
 PT  
 XX  
 PS Disclosure; SEQ ID NO 7828; 1701pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (AB11004-ABA21534) and proteins  
 CC (AB114678-AB18001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 21100 BP; 4855 A; 4720 C; 5213 G; 6312 T; 0 other;  
 Query Match 1.28; Score 19; DB 22; Length 21100;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 689 aacagacagccttggaagc 707  
 Db 5304 AACAGACACCCCTGGAAGGC 5286  
 RESULT 5  
 AAF86431  
 ID AAF86431 standard; DNA; 349980 BP.  
 XX  
 AC AAF86431;  
 XX  
 DT 29-OCT-2001 (first entry)  
 XX  
 DE Pyrococcus abyssi genomic fragment #1.  
 XX  
 KW Hyperthermophilic archaeon; hyperthermophilic protein; ds.

XX Pyrococcus abyssi.  
 QS Key Location/Qualifiers  
 PH misc\_feature 300001...349980  
 FT /\*tag= a  
 FT /note= "This sequence overlaps with the 5' end of  
 FT AAH41223"  
 XX  
 XX FR2792651-A1.  
 XX 27-OCT-2000.  
 XX  
 XX 21-APR-1999; 99FR-0005034.  
 XX 21-APR-1999; 99FR-0005034.  
 XX (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (IFRE-) IFREMER INST FR RECH EXPL MPR.  
 XX  
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 PI Querellou J, Weissbach J, Saurin W, Heilig R;  
 XX WPI; 2001-126236/14.  
 XX  
 XX New nucleotide sequences isolated from Pyrococcus abyssi encode  
 PT proteins useful in industry -  
 XX  
 XX Claim 1: Page 183-279; 1657pp; French.  
 XX  
 XX The present invention relates to the genomic sequence of Pyrococcus  
 CC abyssi and P. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a  
 CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal  
 CC vents. The present sequence is a fragment of the genomic sequence of P.  
 CC abyssi. The 3' end of this sequence overlaps with the 5' end of AAH41223.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as W0200065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAH75903-AAH75920 and AAG66436.  
 XX  
 XX Sequence 349980 BP; 95167 A; 78444 C; 78870 G; 97499 T; 0 other;  
 SQ  
 Query Match 1.2%; Score 19; DB 22; Length 349980;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 837 tggggccttaatttttt 855  
 Db 44109 tggggccttaatttttt 44127  
 RESULT 6  
 ID AAA43311/C  
 XX AAA43311 standard; cDNA; 432 BP.  
 AC AAA43311;  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Xenopus secreted expressed sequence tag SEQ ID NO:2051.  
 XX  
 KW Human; mouse; xenopus; rat; secreted expressed sequence tag; SEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;  
 KW antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW

KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KW tumour; infection; depression; psoriasis; ss.  
 XX Xenopus sp.  
 OS  
 XX W0200021990-A1.  
 PN  
 XX 20-APR-2000.  
 PD  
 XX 15-OCT-1999; 99WO-US24205.  
 PF  
 XX 15-OCT-1998; 98US-0104435.  
 PR  
 XX (GEMV ) GENETICS INST INC.  
 XX  
 XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
 PI Metberg D, Treacy M;  
 PI WPI; 2000-317937/27.  
 DR  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (seSTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 XX Claim 1: Page 589; 618pp; English.  
 PS  
 XX AAA41261 to AAA43419 represent specifically claimed secreted expressed  
 CC sequence tags (seSTs), isolated from human, mouse, xenopus and rat  
 CC tissue sources. The seSTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antitumor; antidiabetic;  
 CC antiasthmatic; vulnary; anticonvulsant; antidepressant; antipsoriatic;  
 CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The seSTs can be used for gene  
 CC therapy and in vaccines. The seSTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the seSTs. Proteins encoded by the seSTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders,  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 432 BP; 144 A; 99 C; 105 G; 84 T; 0 other;  
 Query Match 1.1%; Score 18; DB 21; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 445 ttgctccacaggaagc 462  
 Db 249 TTGCTCCACAGGAAGCAC 232  
 RESULT 7  
 ID AAZ65301 standard; DNA; 1100 BP.  
 XX AAZ65301;  
 AC AAZ65301;  
 XX 23-MAR-2000 (first entry)  
 DT  
 XX





KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
OS Drosophila melanogaster.  
PN WC200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US09231.  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX (PERE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
DR New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX Claim 1; SEQ ID NO 14404; 21bp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (AB57737-AB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX Sequence 1158 BP; 252 A; 285 C; 329 G; 292 T; 0 other;  
SQ

Query Match 1.1%; Score 18; DB 23; Length 1158;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 50 ccacatcctgtgagccg 67  
Db 529 CCATCATCTCTGGAGCCG 512  
|||||  
|||||

RESULT 10  
AAA89186  
ID AAA89186 standard; cDNA; 1494 BP.  
XX AAA89186;  
AC  
XX 19-MAR-2001 (first entry)  
DT  
XX Human oxidoreductase OXRD-2 cDNA clone 1666030CB1.  
DE  
XX OXRD-2; human; oxidoreductase; cell proliferation;  
KW neurological disease; smooth muscle disease; autoimmune disease;  
KW inflammation; antiproliferative; neuroprotective;  
KW immunosuppressive; antiinflammatory; gene therapy; diagnosis; ss.  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 29..1318  
FT CDS /\*tag= a  
FT  
XX WC200071679-A2.  
PN  
XX

PD 30-NOV-2000.  
XX 19-MAY-2000; 2000WO-US13879.  
XX 20-MAY-1999; 99US-0135049.  
PR 27-MAY-1999; 99US-0136740.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Tang YT, Baughn MR, Lu DAM;  
PI WPI; 2001-025146/03.  
DR P-PSDB; AAB19927.  
XX New human oxidoreductase proteins useful for diagnosing, treating or  
XX preventing proliferative, neurological, genetic, smooth muscle,  
XX autoimmune or inflammatory disorders associated with abnormal  
XX expression of oxidoreductase proteins -  
XX Claim 5; Page 89; 95pp; English.  
XX The present sequence is that of Incyte clone 1666030CB1 encoding  
CC human oxidoreductase OXRD-2 (see AAB19927). The gene maps to  
CC chromosome 1 between 278.8 and 292.7 cM. The clone was isolated  
CC from human breast tissue cDNA library BRSTNOT09. The encoded  
CC protein shows homology to Plasmodium falciparum FAD-dependent  
CC oxidoreductase. OXRD-2 is expressed in nervous, reproductive and  
CC gastrointestinal tissue, and may be involved in cell proliferation  
CC and inflammation. The invention provides OXRD-1 to -8 polypeptides  
CC (see AAB19326-33) and polynucleotides (see AAA89185-92). It also  
CC provides methods for using these polypeptides and polynucleotides  
CC for diagnosing, treating or preventing disorders associated with  
CC expression of OXRD, especially cell proliferative, neurological,  
CC genetic, smooth muscle, and autoimmune/inflammatory disorders.  
CC A vector encoding OXRD or its fragments is useful for somatic or  
CC germ-line gene therapy for treating such disorders. The  
CC polynucleotides may also be used to monitor regulation of OXRD  
CC levels during therapeutic intervention. A DNA fragment comprising  
CC nucleotides 173-217 of the present sequence can be used as a  
XX hybridization probe.  
SQ Sequence 1494 BP; 419 A; 301 C; 368 G; 406 T; 0 other;  
Query Match 1.1%; Score 18; DB 22; Length 1494;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1586 aagagaaggtggccaatt 1603  
Db 716 aagagaaggtggccaatt 733  
|||||  
|||||

RESULT 11  
AAK94526  
ID AAK94526 standard; cDNA; 1549 BP.  
XX AAK94526;  
AC  
XX 06-NOV-2001 (first entry)  
DT  
XX Human full-length cDNA, SEQ ID NO: 3398.  
DE  
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.  
KW Homo sapiens.  
OS  
XX EP1130094-A2.  
PN  
XX 05-SEP-2001.  
PD  
XX 07-JUL-2000; 2000EP-0114089.  
PF  
XX

PR 08-JUL-1999; 99JP-0194486.  
 PR 11-JAN-2000; 2000JP-0118774.  
 PR 02-MAY-2000; 2000JP-0183765.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX WPI: 2001-524255/58.  
 DR P-PSDB; AAM93594.  
 DR  
 PT 830 Primers useful for synthesizing full length cDNA clones and their  
 XX use in genetic manipulation.  
 PT  
 PS Claim 8; SEQ ID NO 3399; 1380pp + sequence listing: English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been  
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
 CC molecules have been determined. Primers for synthesizing the full length  
 CC cDNA are useful for clarifying the function of the protein encoded by  
 CC the cDNA. The full length clones were obtained by construction of full  
 CC length enriched cDNA libraries that were synthesised by the oligo-capping  
 CC method. The primers enable the production of the full length cDNA easily  
 CC without any special methods. The present sequence is a full length  
 CC human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in CD-ROM format directly from EPO.  
 CC  
 XX Sequence 1549 BP; 408 A; 326 C; 400 G; 415 T; 0 other;  
 SQ

Query Match 1.18; Score 18; DB 22; Length 1549;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1586 aagagaaggtggccaatt 1603  
 ||||||||||||||||  
 Db 791 aagagaaggtggccaatt 808

RESULT 12  
 ABL15755  
 ID ABL15755 standard; cDNA: 2095 BP.  
 AC ABL15755;  
 XX  
 DT 26-MAR-2002 (first entry)  
 DE  
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 41747.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX  
 PF 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PR  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI: 2001-656860/75.  
 DR P-PSDB; ABB71652.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Claim 1; SEQ ID NO 41747; 21pp + Sequence Listing: English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pat\_sequences.  
 CC  
 XX Sequence 2095 BP; 607 A; 520 C; 517 G; 451 T; 0 other;  
 SQ

Query Match 1.18; Score 18; DB 23; Length 2095;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 tcaggaggtggccaagaa 611  
 ||||||||||||||||  
 Db 1218 tcaggaggtggccaagaa 1235

RESULT 13  
 AAC76642  
 ID AAC76642 standard; cDNA: 2171 BP.  
 XX  
 AC AAC76642;  
 XX  
 DT 08-FEB-2001 (first entry)  
 DE  
 XX Human ORFX ORF2197 polynucleotide sequence SEQ ID NO:4393.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antineumatic; antithyroid;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW severe combined immunodeficiency; systemic lupus erythematosus; infection;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 XX 31-MAR-2000; 2000WO-US08621.  
 XX  
 PF 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 PR  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Shimkets RA, Leach M;  
 PI WPI: 2000-602362/57.  
 DR

DR P-PSDB; AAB42433.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 5; Page 3582-3583; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC osteoplastic; antiparkinsonian; neurotropic; neuroprotective;  
 CC osteoplastic; anticonvulsant; antithrombotic; immunosuppressive;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antididiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 CC antithyroid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX  
 SQ Sequence 2171 BP; 584 A; 429 C; 505 G; 653 T; 0 other;

Query Match 1.1%; Score 18; DB 21; Length 2171;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1586 aagagaagtggtgccaatt 1603  
 DB 727 aagagaagtggtgccaatt 744  
 |||||

RESULT 14  
 AAH17299/c  
 ID AAH17299 standard; cDNA; 2624 BP.  
 XX  
 AC AAH17299;  
 XX  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:16700.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX Homo sapiens.  
 OS  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX

DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 16700; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC polynucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 2624 BP; 514 A; 623 C; 710 G; 777 T; 0 other;

Query Match 1.1%; Score 18; DB 22; Length 2624;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1530 agcaacacgctgctttaa 1547  
 DB 1214 AGCAACACGCTGCTTAAAC 1197  
 |||||

RESULT 15  
 ABL20976  
 ID ABL20976 standard; DNA; 3158 BP.  
 XX  
 AC ABL20976;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 14401.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 00:46:03 ; Search time 7820.58 Seconds  
(without alignments)  
4383.008 Million cell updates/sec

Title: US-09-882-694A-2

Perfect score: 1638

Sequence: 1 atctgcgcaccagcaactc.....tgcttcgtctccctgtaa 1638

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.vi.\*
- 14: gb.vl.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vl.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htgo.inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1638	100.0	1638	6	AX076843
2	1184	72.3	1691	6	AX076842
3	22	1.3	6047	1	ECIS1294
4	22	1.3	8561	1	EC008006
5	20	1.2	29716	2	AC103823
6	20	1.2	63585	2	AC026983
7	20	1.2	99015	2	AC091870
8	20	1.2	149179	9	AC021694
9	20	1.2	158781	9	AC024978
10	20	1.2	160099	9	AC067844
11	20	1.2	162054	2	AC046201
12	20	1.2	162515	2	AC091097
13	20	1.2	163452	2	AC079184
14	20	1.2	164436	2	AC051656
15	20	1.2	192974	9	AC087565
16	20	1.2	196501	9	AC005908
17	20	1.2	202993	2	AC060775
18	20	1.2	208361	2	AC104343
19	20	1.2	210617	2	AC034201
20	19	1.2	2011	10	AF021802
21	19	1.2	4195	10	MUSAGOUA
22	19	1.2	4718	10	MUSAGOUA
23	19	1.2	7040	8	MZETRACTA
24	19	1.2	10033	2	AC109682
25	19	1.2	12911	1	AE005242
26	19	1.2	42140	9	AC087234
27	19	1.2	52783	2	AC100403
28	19	1.2	63609	9	AL133343
29	19	1.2	67175	9	AL392183
30	19	1.2	69599	2	AC102128
31	19	1.2	103681	9	HS568C11
32	19	1.2	110000	2	AC092921_0
33	19	1.2	110000	2	AC105469_1
34	19	1.2	116557	9	AC073090
35	19	1.2	134682	2	AC092553
36	19	1.2	141274	2	AF374376
37	19	1.2	147207	2	AC078890
38	19	1.2	153087	2	AC017072
39	19	1.2	161105	2	AC024718
40	19	1.2	169934	2	AC021870
41	19	1.2	173748	2	AC024553
42	19	1.2	176530	2	AC099361
43	19	1.2	176982	9	AC092846
44	19	1.2	192389	9	AC007182
45	19	1.2	198050	1	AL646061

ALIGNMENTS

RESULT 1	AX076843	Sequence 2 from Patent WO0105980.	1638 bp	DNA	linear	PAT 22-FEB-2001
LOCUS	AX076843					
DEFINITION	AX076843					
ACCESSION	AX076843.1	GI:13121518				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
	Exophiala spinifera					
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothyrionycetes;					
	Chaetothyrionales; Herpotrichiellaceae; Exophiala.					
	Herpotrichiellaceae; Exophiala.					
REFERENCE						
AUTHORS	Duvick, J.P., Maddox, J., Gilliam, J., Folkerts, O. and Crasta, O.R.					
TITLE	Compositions and methods for fumonisin detoxification					
JOURNAL	Patent: WO 0105980-A 2 25-JAN-2001;					
	Pioneer Hi-Bred International Inc. (US) ; Curagen Corporation (US)					
FEATURES						
	Location/Qualifiers					
	1..1638					
	/organism="Exophiala spinifera"					
	/db_xref="taxon:91928"					
	/note="flavin monooxygenase, coding sequence"					
	388 a 432 c 462 g 356 t					
BASE COUNT						

## ORIGIN

Query Match 100.0%; Score 1638; DB 6; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtcgccaccagcaatccagagggatgttccctgcgcagcgcacatcggtt 60  
Db 1 ATGTGCGCCACCAGCAATCCAGAGGGGATTTCCGTCGCATCGCAGCGCATCATCGTT 60  
Qy 61 ggaagcgccctcagcgcatctcgtgtgtacaaaattgcaaaagctcaagctcaagcc 120  
Db 61 GGAGCGCGGCTCAGCGGCATCTGCTGTGTACAAAATTGCAAAAGCTCAACACTCAAGCC 120  
Qy 121 aaaaatttcgagggagcccccagatttggcggcgtctggaactggaaccctaccctggc 180  
Db 121 AAAATCTTCGAGGGAGCCCCGATTTTGGCGGCTCTGGCACTTGAACCCGCTACCCCTGGC 180  
Qy 181 gctcgtgttgatcggagagcgccttctacaaactgaacattcccgaagcatgaaagac 240  
Db 181 GCTCGTGTGATTTCGAGAGCGCCCTTACCAACTGAACATTCGCCGAAGTATGGAAGAC 240  
Qy 241 tgaactgtctgscgtatctcgaacagaaagagttgctatgtatgttaccactgt 300  
Db 241 TGAACCTGGTCTTGCGGCTATCTCTACCCAGAAAGAGTTGCTGTATGTTCACCACTGT 300  
Qy 301 gacaagatccggggcttgagaaagacgtctacttcggagctgaggtgggtgtgatgcgcg 360  
Db 301 GACAAGATCCGGGGCTTGAGAAAAGACGCTTACTTCGGAGCTGAGGTGGTTGATGCCGG 360  
Qy 361 tatgccagatctgggacactcgactctcaagcgtcggtcggtgcgcgtatgttgcagcgca 420  
Db 361 TATGCCAGAGATCTGGGCACCTGGACTGTCAAGACGTGGCTGGCCATGTGTGCAGCGCA 420  
Qy 421 aagtatctatctcgtctacggggtgtctccaaaggagcaacactcccgcactcccgcgc 480  
Db 421 AAGTATCTCATTTCTGCTACGGGTGTCTCCACAGGAAGCACACTCCCGCACTCCCGGCG 480  
Qy 481 ctgcgcatttcaagggaggtgattcttcagtgagtcctggcagcaagacttcgacgca 540  
Db 481 CTGCGCGATTCAACGGGAAGTGATTCATTCGAGTGCCTGGCACGAAGACTTCGACGCA 540  
Qy 541 gaggccagagatccgctcagtcagtcggtcggtcggtcggtcggtcggtcggtcggtcggt 600  
Db 541 GAGGGCCAGAGAGTCCCGCTCATCGTCCCGGGGCTCAACAGCATCCAGATGTGTTCAGGAG 600  
Qy 601 ttggccaaaggtgaccaggtaccatgtttatgcaaggcggagctattgtctgccc 660  
Db 601 TTGGCCAAAGAGCTGACAGGTAACTACCATGTTTATGGAAGGGCGGAGCTATTGTCTGCC 660  
Qy 661 atcgcccaacgacgatgataaggaacagacagcagcctggagggcctactaccaccag 720  
Db 661 ATCGGCAACGAACGATGATAGGAACCAACAGACAGCTTGAAGGCTACTACCCACG 720  
Qy 721 ctgtttgaagcaggtcgaaagctcgtgattggattccggtccagggcagcgtcggttggc 780  
Db 721 CTGTTTGAAGCAGTCTGAAGTCTCGGATTGGATTCCCGGTCCAGGACCGCTCGGTGGC 780  
Qy 781 atctttgaagtcagcccgagcagcggtgagccctatttcgaagagttgtgggagcgtggg 840  
Db 781 ATCTTTGAAGTCAAGCGCGGAGGCGCTATTTCGAAGAGTTGTGGGAGCGGTGGG 840  
Qy 841 gcttttaatttcttgcttcgagtcaggaagatcatgttatcaaaaagggcaaccga 900  
Db 841 GCCTTTAATTTCTGCTTGCAGTACCGAAGAGTATGTTGACAAAAGGCAACCGA 900  
Qy 901 ctggtctatgactcttgggcaaaaagactcgatctcgtatcgttcactcggaagaga 960  
Db 901 CTGGTCTATGACTTCTGGSCAAAAGACTCGATCTCGTATCGTCAATCCGCAAGAGA 960  
Qy 961 gatctaatgctcctctgagcgccgtactgtgttcggtaccagcgctcccactggag 1020  
Db 961 GATCTCATGCTCTCTCGAGCGCGCTACTGTTCCGTTACCAAGCGCTCCCCACTGGAG 1020

Db 961 GATCTCATGCTCTCTCGAGCGCGCTACTGTTCCGTTACCAAGCGCTCCCCACTGGAG 1020  
Qy 1021 agcactactcaaaatctgagcaagccgagcgtcgaaattgtgaattagaaacaatcg 1080  
Db 1021 ACCACTACTACGAATGCTGGCAAGCCGCGCTCGAAATTTGTAATCTAGAACAACTCG 1080  
Qy 1081 ccaattgtgctgtctacaagacaggtgtgtcttctgtgagtgagcggcagcaagggaatgc 1140  
Db 1081 CCCATTGTGCTGTTCACAAAGACAGGTTGCTCTTGTAGTGAGCGCAAGAGGAATGC 1140  
Qy 1141 gacacgactcgtgtggcagcggttttcgacagtttcaactggctcattgacacatatgggc 1200  
Db 1141 GACACGATCTGCTGGCGACGGGTTTCGACAGTTTCACCTGGCTCATTCACACATATGGC 1200  
Qy 1201 ttgaaaaaagacagcgagtgagacctgaaggagggtgtggaagaatgagcatatctacttat 1260  
Db 1201 TTGAAAACAAAGCAGCGAGTGGACCTGAAGAGAGGTGTGAAAAGATGGCATATCTACTTAT 1260  
Qy 1261 atggagctctctctcatggcttcccccgaatgcttcttcgtcggccacggtccaagcccg 1320  
Db 1261 ATGGAGTCTCTCTCATGCTTCCCAATGCTTCTTGTCCGACGCTCAAGCCCG 1320  
Qy 1321 accgtcttcccaacgcccacagcatcatagaaacccaagtcgactgtatgcgcgataca 1380  
Db 1321 ACCGTCTTCTCCAAACGCCCCACGATCATAGAAACCAAGTCGACTTGTATGCCCGATACA 1380  
Qy 1381 attgcaagttgagggccgagcagccacgtccgtttgaggcgacgaaatcagcaagaag 1440  
Db 1381 ATTGCAAGTTGAGGCGGACGACGCGCTGCTGTTGAGGCGACGAAATACGACAAAGAG 1440  
Qy 1441 gcatggtcgatgatgttgcgaagatgaacgagcagcactcttctcccttgacggattcg 1500  
Db 1441 GCATGGTCTGATTTATGATGGCCAAAGATGAACGAGCACACTCTCTTCCCTTGACGGATTTCG 1500  
Qy 1501 tggctgactgagggcaacatccctggggaagaacacgtgcttttaacattcattaggggg 1560  
Db 1501 TGGTGGACTGGAGGCAACATCCCTGGGAAAGCAACAGTGTCTTAACTTTCATAGGCGGG 1560  
Qy 1561 attgctctctatgagcagatctgtcagagaggggtggccaattggagtggtttgatgtg 1620  
Db 1561 ATTGCTCTCTATGAGCAGATCTGTCAAGAGAAAGGTGGCCAATTGGATGGGTTTGATGTG 1620  
Qy 1621 ctctcatgctccctgctaa 1638  
Db 1621 CTTCATGCTCCCTGCTAA 1638  
RESULT 2  
AX076842  
LOCUS 1691 bp DNA linear PAT 22-FEB-2001  
DEFINITION Sequence 1 from Patent WO0105980.  
ACCESSION AX076842  
VERSION AX076842.1 GI:13121517  
KEYWORDS  
SOURCE Exophiala spinifera.  
ORGANISM Exophiala spinifera  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothryiomycetes;  
Chaetothryiales; Herpotrichiellaceae; mitosporic  
Herpotrichiellaceae; Exophiala.  
REFERENCE 1 (bases 1 to 1691)  
AUTHORS Duvick, J. P., Maddox, J., Gilliam, J., Folkerts, O., and Crasta, O. R.  
TITLE Compositions and methods for fumonisin detoxification  
JOURNAL Patent: WO 0105980-A 1 25-JAN-2001;  
Pioneer Hi-Bred International Inc. (US) ; Curagen Corporation (US)  
FEATURES  
source Location/Qualifiers  
1..1691  
/organism="Exophiala spinifera"  
/db\_xref="taxon:91928"  
/note="flavin monooxygenase with intron"  
BASE COUNT 397 a 444 c 480 g 370 t  
ORIGIN





```

Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 842 ccttaattttttgttgccca 863
      |||||
Db 362 CCTTTAAATTTCTGCTGCCA 341

RESULT 4
LOCUS ECO08006/c 8561 bp DNA circular BCT 25-APR-2000
DEFINITION Escherichia coli kanamycin resistance plasmid PUB2380.
ACCESSION AJ008006
VERSION AJ008006.1 GI:3413194
KEYWORDS aminoglycoside 3'-phosphotransferase; aph; mob1 gene; mob2 gene;
mob3 gene; mob4 gene; mob6 gene; mob7 gene; mob8 gene; tnp1294
gene; transposase.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
          Escherichia.
REFERENCE 1 (bases 1 to 8561)
AUTHORS Albiger,B., Bennett,P.M., Commanducci,A., Dodd,H.M. and Lett,M.C.
TITLE PUB2380 : a Cold-like resistance plasmid with three replication
        origins for three different functions
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8561)
AUTHORS Bruyns,E.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1998) Bruyns E., Institute for Immunology,
        University of Heidelberg, Im Neuenheimer Feld 305, 69120
        Heidelberg, GERMANY
COMMENT Related sequence: X82430.
FEATURES             Location/Qualifiers
     source           1..8561
                     /organism="Escherichia coli"
                     /plasmid="PUB2380"
                     /strain="SE53"
                     /db_xref="taxon:562"
                     /map="0 to 8.5"
     misc_feature     1..619
                     /note="ColE1-like replication region"
                     /evidence=experimental
     -35_signal       8..14
                     /note="RNAII"
                     /evidence=experimental
     -10_signal       32..34
                     /note="RNAII"
                     /evidence=experimental
     misc_feature     46
                     /note="RNAII transcription start"
                     /complement(149)
     misc_feature     149
                     /note="RNAI transcription start"
                     /complement(155..161)
     -10_signal       155..161
                     /note="RNAI"
     -35_signal       complement(178..184)
                     /note="RNAI"
     rep_origin       619
                     /note="ColE1-like origin of replication"
                     /direction=RIGHT
     misc_feature     complement(828..829)
                     /note="ColE1-like relaxation (nic) cut site"
     gene             972..1109
                     /gene="mob1"
     CDS              972..1109
                     /gene="mob1"
                     /codon_start=1
                     /transl_table=11
                     /product="putative mob1 protein"
                     /protein_id="CAA07826.1"
                     /db_xref="GI:3413195"
                     /db_xref="SPTREMBL:Q57132"
                     /translation="MSEVFVTPCEAAAVKLGVVVKRFTNIGLFCVQFVELQQ
                     R"
1206..1553
/gene="mob2"
1206..1553
/gene="mob2"
/codon_start=1
/transl_table=11
/product="putative mob2 protein"
/protein_id="CAA07827.1"
/db_xref="GI:3413196"
/db_xref="SPTREMBL:Q51618"
/translation="MTPKMERMLTIRVTDDEHARLERCEGKQLAVMMRRVCLGPV
ARSGKLPTLAPLLRLAALAAIGNNLTARKVNSGOWSSGDRVQVVAALMAIGDELRL
RLAVREQGARDDS"
1543..3096
/gene="mob3"
1543..3096
/gene="mob3"
/codon_start=1
/transl_table=11
/product="putative mob3 protein"
/protein_id="CAA07828.1"
/db_xref="GI:3413197"
/db_xref="SPTREMBL:Q51619"
/translation="MIVKFKAGKGGCGPDVLLGRNRREGATVLOGIPREVRELI
DATPPAKKVTSGVLSFAEKELPGREKYNASPEVLMPELKNQYSILAVSHQDKGR
LELNFVPMELQTKRRLQFYIDRADNPIDAWQTLNHHYGLHDNPANENRRTLTLPL
DNLPTKQALAEQVTRGIDALYHAGEIKRQDYIQALTEAGLEVVRVTRKSSISADPN
GGKNTLKAGFYEQSPADGKGVREKRESRIYRENAEQRVQARRICRGCDIKHDE
NQRYSPVHSLDRGIAGKTPGRGERGDDAAQEGRVKAGREYGHDTGDSLSVYREMR
DALVSWEDTGEPRNQEAGRTAETREDMGRCVACGREQIIPCFSVREISGGDSLS
GERVGTSEGVTSDRAGNTFAERLAAATGLYAAARMGERLGRGAEDVFAYATGORD
AERAGHAVESAGATLERADRTLEPVIQRELEIRERLQEREHVLSLEREROPEIOER
TLDGPSLGH"
1559..1762
/gene="mob4"
1559..1762
/gene="mob4"
/codon_start=1
/transl_table=11
/product="putative mob4 protein"
/protein_id="CAA07829.1"
/db_xref="GI:3413198"
/db_xref="SPTREMBL:O86201"
/translation="MPGEKVAVVRLITTCGGSVTAKAQCPRVYRKKGSGNSMPPHL
BENTRPVFCRSRRRSRRREDVKK"
2231..2749
/gene="mob6"
2231..2749
/gene="mob6"
/codon_start=1
/transl_table=11
/product="putative mob6 protein"
/protein_id="CAA07830.1"
/db_xref="GI:3413199"
/translation="MSNLQGTGAEFKKLKERAEKTEKMLNNEFRRLGESVSEAVTNS
ETKIRDAIALFTASTEESLEKHKREGVKAMQHRDVLKLAGNTGMMLGIVFLFLFTA
SGGTLWYLGGRIOANLEIRKQETLQKLNAKTGWGVFVODGNRKFVLVLPYKGSAPVI
PFQKGKWHVHKE"
2813..2989
/gene="mob7"
2813..2989
/gene="mob7"
/codon_start=1
/transl_table=11
/product="putative mob7 protein"
/protein_id="CAA07831.1"
/db_xref="GI:3413200"
/translation="MORLSEWESAFVELOKMFSLTQRDNAMLNERVMQISOQVQHLSE
QTERLSQLYSENWR"
3074..3181
/gene="mob8"
3074..3181
/gene="mob8"

```

```

/codon_start=1
/transl_table=11
/product="putative mob8 protein"
/protein_id="CAA07832.1"
/db_xref="GI:3413201"
/db_xref="SPTREMBL:O86202"
/translation="MALRWDETGMKSGFSLGSGNVACTLWLHRYSVG"
3470..4522
/gene="ORF1"
3470..4522
/gene="ORF1"
/transl_table=1
/codon_start=1
/evidence-experimental
/product="hypothetical protein"
/protein_id="CAA07833.1"
/db_xref="GI:3413202"
/db_xref="SPTREMBL:O86203"
/translation="MNLQORINKLPQLSSSFCKIDINHSFIENETSKDKIEDLLAR
KWSGQPCVFGKASKKIKGLDFHLSIVNSFQLNDGHLFDELRLNRVRFKERRR
GEVSAHLIFIHPLGARFSEELVDIOKYLCSHMPECYFIKEDGVITYTESVPPQDRD
GLKIRAGVNFYSSAHRTRNDRRIPGILISVNAFMRZLAERGFYKDOQALA
DIKMTIQSVGGKGYSHPEGISTTWHSNLDKRGCPVHTGSSYYSGFYHTDYLIFG
ELTKDLRLHREIENSDDPMIFNNVLFYVSLEEFIDDPYGEFTGIPVDQASMPFNSF
QPKRFENKPLYEKEDD"
4515..5390
/gene="ORF2"
4515..5390
/gene="ORF2"
/transl_table=1
/codon_start=1
/evidence-experimental
/product="hypothetical protein"
/protein_id="CAA07834.1"
/db_xref="GI:3413203"
/db_xref="SPTREMBL:O86204"
/translation="MNMIMKMLFAFLAFAWSGAIVMSKGVBELDAGGVFFWQICS
AVILSVLFIKRLKPLNSKAILAYATGVPEPLAYFTLYGLYFISAGSASLIPLS
ESVFLVLSVLCVSKINSQYFIPLFGAIVGSLVALPOLNSGNNVIGYCLVAG
VLSAAVYVLSRLKISFEPITLLSGQUTFSLLSLFIFLSKSIYLPUSVLIIFS
SGILQVFLAFPCFYLNLPLVHLGIMLYFIPVLAFLSLWFLDEQFTFLQGVGFV
TVLSVFLNKKYSEE"
5401..7088
/organism="Escherichia coli"
/transposon="IS1294"
/db_xref="taxon:562"
5401
/feature_start=IS1294"
/evidence-experimental
/complement(5715..6770)
/gene="tnp1294"
/complement(5715..6770)
/gene="tnp1294"
/codon_start=1
/transl_table=11
/evidence-experimental
/product="transposase"
/protein_id="CAA07835.1"
/db_xref="GI:3413204"
/db_xref="SPTREMBL:Q47307"
/translation="MLSGFTPLKRLFTANOCWTSFLDAGGLRDIEEAVTKMLACR
TRILGVKEFCDDNPDCHVKYLTNCSGRACPCGKATDLWTATQLNRLPDCDWHL
VFPLPDLVWPFESNLLNCDRLAVENLLYAARKGLEGICFAIHTYGRRLNHP.

Query Match 1.3%; Score 22; DB 1; Length 8561;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 cctttaattttcttgctgcc 863
|||||
DB 3663 CCTTTAAATTTCTTGCTGCC 3642

```

```

RESULT 5
AC103823/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

AC103823 Homo sapiens chromosome 8 clone CTD-2121K3 map 8, LOW-PASS SEQUENCE SAMPLING.
AC103823.1 GI:17149708
HTG: HTGS_PHASE0.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone CTD-2121K3
Unpublished
2 (bases 1 to 29716)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N., Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,B., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21811
Center clone name: 2121_K_3
-----
* NOTE: This record contains 36 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 699: contig of 699 bp in length
700 799: gap of 100 bp
800 1502: contig of 703 bp in length
1503 1602: gap of 100 bp
1603 2296: contig of 694 bp in length
2297 2396: gap of 100 bp
2397 3101: contig of 705 bp in length
3102 3201: gap of 100 bp

```

```

3202 3930: contig of 729 bp in length
3931 4030: gap of 100 bp
4031 4765: contig of 735 bp in length
4766 4865: gap of 100 bp
4866 5557: contig of 692 bp in length
5558 5657: gap of 100 bp
5658 6440: contig of 783 bp in length
6441 6540: gap of 100 bp
6541 7282: contig of 742 bp in length
7283 7382: gap of 100 bp
7383 8131: contig of 749 bp in length
8132 8231: gap of 100 bp
8232 8936: contig of 705 bp in length
8937 9036: gap of 100 bp
9037 9697: contig of 661 bp in length
9698 9797: gap of 100 bp
9798 10478: contig of 681 bp in length
10479 10578: gap of 100 bp
10579 11317: contig of 739 bp in length
11318 11417: gap of 100 bp
11418 12153: contig of 736 bp in length
12154 12253: gap of 100 bp
12254 12985: contig of 732 bp in length
12986 13085: gap of 100 bp
13086 13836: contig of 751 bp in length
13837 13936: gap of 100 bp
13937 14715: contig of 779 bp in length
14716 14815: gap of 100 bp
14816 15544: contig of 729 bp in length
15545 15644: gap of 100 bp
15645 16430: contig of 786 bp in length
16431 16530: gap of 100 bp
16531 17266: contig of 736 bp in length
17267 17366: gap of 100 bp
17367 18091: contig of 725 bp in length
18092 18191: gap of 100 bp
18192 18874: contig of 683 bp in length
18875 18974: gap of 100 bp
18975 19708: contig of 734 bp in length
19709 19808: gap of 100 bp
19809 20545: contig of 737 bp in length
20546 20645: gap of 100 bp
20646 21363: contig of 718 bp in length
21364 21463: gap of 100 bp
21464 22193: contig of 730 bp in length
22194 22293: gap of 100 bp
22294 23021: contig of 728 bp in length
23022 23121: gap of 100 bp
23122 23795: contig of 674 bp in length
23796 23895: gap of 100 bp
23896 24630: contig of 735 bp in length
24631 24730: gap of 100 bp
24731 25496: contig of 766 bp in length
25497 25596: gap of 100 bp
25597 26378: contig of 782 bp in length
26379 26478: gap of 100 bp
26479 27161: contig of 683 bp in length
27162 27261: gap of 100 bp
27262 27979: contig of 718 bp in length
27980 28079: gap of 100 bp
28080 28836: contig of 757 bp in length
28837 28936: gap of 100 bp
28937 29716: contig of 780 bp in length.

```

FEATURES

Source

```

1. 29716
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="8"
   /map="8"
   /clone="CTD-2121K3"
   /clone_lib="CITD Human BAC"
7625 a 5404 c 5170 g 7924 t 3593 others

```

BASE COUNT  
ORIGIN

```

Query Match          1.2%; Score 20; DB 2: Length 29716;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 ccttaattttcttcttc 861
Db 13756 CCTTAATTTTCTTCTTC 13737

RESULT 6
AC026983/c
LOCUS
DEFINITION Homo sapiens chromosome 8 clone RP11-376P3 map 8, LOW-PASS SEQUENCE
SAMPLING.
AC026983
VERSION AC026983.1 GI:7328861
KEYWORDS HTG: HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 63585)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-376P3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 63585)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,R., Lacombe,K., Lamazares,R., Landers,T., Lehotzky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meidrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Scheuer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WTHR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8673
Center clone name: 376_P_3
-----
* NOTE: This record contains 66 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that

```

\* the record is updated, the accession number will  
\* be preserved.

\* 839 938: gap of 100 bp  
\* 839 1825: contig of 838 bp in length  
\* 1826 1925: gap of 100 bp  
\* 1926 2786: contig of 861 bp in length  
\* 2787 2886: gap of 100 bp  
\* 2887 3770: contig of 884 bp in length  
\* 3771 3870: gap of 100 bp  
\* 3871 4699: contig of 829 bp in length  
\* 4700 4799: gap of 100 bp  
\* 4800 5665: contig of 866 bp in length  
\* 5666 5765: gap of 100 bp  
\* 5766 6632: contig of 867 bp in length  
\* 6633 6732: gap of 100 bp  
\* 6733 7595: contig of 863 bp in length  
\* 7596 7695: gap of 100 bp  
\* 7696 8550: contig of 855 bp in length  
\* 8551 8650: gap of 100 bp  
\* 8651 9524: contig of 874 bp in length  
\* 9525 9624: gap of 100 bp  
\* 9625 10513: contig of 889 bp in length  
\* 10514 10613: gap of 100 bp  
\* 10614 11471: contig of 858 bp in length  
\* 11472 11571: gap of 100 bp  
\* 11572 12465: contig of 894 bp in length  
\* 12466 12565: gap of 100 bp  
\* 12566 13397: contig of 832 bp in length  
\* 13398 13497: gap of 100 bp  
\* 13498 14353: contig of 856 bp in length  
\* 14354 14453: gap of 100 bp  
\* 14454 15274: contig of 821 bp in length  
\* 15275 15374: gap of 100 bp  
\* 15375 16246: contig of 872 bp in length  
\* 16247 16346: gap of 100 bp  
\* 16347 17201: contig of 855 bp in length  
\* 17202 17301: gap of 100 bp  
\* 17302 18189: contig of 888 bp in length  
\* 18190 18289: gap of 100 bp  
\* 18290 19181: contig of 892 bp in length  
\* 19182 19281: gap of 100 bp  
\* 19282 20154: contig of 873 bp in length  
\* 20155 20254: gap of 100 bp  
\* 20255 21084: contig of 830 bp in length  
\* 21085 21184: gap of 100 bp  
\* 21185 22051: contig of 867 bp in length  
\* 22052 22151: gap of 100 bp  
\* 22152 22983: contig of 832 bp in length  
\* 22984 23083: gap of 100 bp  
\* 23084 23957: contig of 874 bp in length  
\* 23958 24057: gap of 100 bp  
\* 24058 24948: contig of 891 bp in length  
\* 24949 25048: gap of 100 bp  
\* 25049 25936: contig of 888 bp in length  
\* 25937 26036: gap of 100 bp  
\* 26037 26919: contig of 883 bp in length  
\* 26920 27019: gap of 100 bp  
\* 27020 27904: contig of 885 bp in length  
\* 27905 28004: gap of 100 bp  
\* 28005 28888: contig of 884 bp in length  
\* 28889 29888: gap of 100 bp  
\* 29890 29864: contig of 876 bp in length  
\* 29865 29964: gap of 100 bp  
\* 29965 30817: contig of 853 bp in length  
\* 30818 30917: gap of 100 bp  
\* 30919 31801: contig of 884 bp in length  
\* 31802 31901: gap of 100 bp  
\* 31902 32742: contig of 841 bp in length  
\* 32743 32842: gap of 100 bp  
\* 32843 33744: contig of 902 bp in length  
\* 33745 33844: gap of 100 bp  
\* 33845 34724: contig of 880 bp in length

\* 34725 34824: gap of 100 bp  
\* 34825 35695: contig of 871 bp in length  
\* 35696 35795: gap of 100 bp  
\* 35796 36656: contig of 861 bp in length  
\* 36657 36756: gap of 100 bp  
\* 36757 37649: contig of 893 bp in length  
\* 37650 37749: gap of 100 bp  
\* 37750 38571: contig of 822 bp in length  
\* 38572 38671: gap of 100 bp  
\* 38672 39508: contig of 837 bp in length  
\* 39509 39608: gap of 100 bp  
\* 39609 40471: contig of 863 bp in length  
\* 40472 40571: gap of 100 bp  
\* 40572 41395: contig of 824 bp in length  
\* 41396 41495: gap of 100 bp  
\* 41496 42354: contig of 859 bp in length  
\* 42355 42454: gap of 100 bp  
\* 42455 43353: contig of 899 bp in length  
\* 43354 43453: gap of 100 bp  
\* 43454 44316: contig of 863 bp in length  
\* 44317 44416: gap of 100 bp  
\* 44417 45274: contig of 858 bp in length  
\* 45275 45374: gap of 100 bp  
\* 45375 46244: contig of 870 bp in length  
\* 46245 46344: gap of 100 bp  
\* 46345 47209: contig of 865 bp in length  
\* 47210 47309: gap of 100 bp  
\* 47310 48162: contig of 853 bp in length  
\* 48163 48262: gap of 100 bp  
\* 48263 49117: contig of 855 bp in length  
\* 49118 49217: gap of 100 bp  
\* 49218 50059: contig of 842 bp in length  
\* 50060 50159: gap of 100 bp  
\* 50160 51013: contig of 854 bp in length  
\* 51014 51113: gap of 100 bp  
\* 51114 51972: contig of 859 bp in length  
\* 51973 52072: gap of 100 bp  
\* 52073 52967: contig of 895 bp in length  
\* 52968 53067: gap of 100 bp  
\* 53068 53928: contig of 861 bp in length  
\* 53929 54028: gap of 100 bp  
\* 54029 54878: contig of 850 bp in length  
\* 54879 54978: gap of 100 bp  
\* 54979 55852: contig of 874 bp in length  
\* 55853 55952: gap of 100 bp  
\* 55953 56836: contig of 884 bp in length  
\* 56837 56936: gap of 100 bp  
\* 56937 57771: contig of 835 bp in length  
\* 57772 57871: gap of 100 bp  
\* 57872 58718: contig of 847 bp in length  
\* 58719 58818: gap of 100 bp  
\* 58819 59707: contig of 889 bp in length  
\* 59708 59807: gap of 100 bp  
\* 59808 60663: contig of 856 bp in length  
\* 60664 60763: gap of 100 bp  
\* 60764 61629: contig of 866 bp in length  
\* 61630 61729: gap of 100 bp  
\* 61730 62601: contig of 872 bp in length  
\* 62602 62701: gap of 100 bp  
\* 62702 63585: contig of 884 bp in length.

FEATURES  
source

Location/Qualifiers  
1. .63585  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-376P3"

Query Match 1.2%; Score 20; DB 2; Length 63585;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1279 ggcttcccaatgcctctt 1298

Db 36649 GCGTCCCAATGCTCTCT 36630  
 |||  
 |||

## RESULT 7

AC091870/c 99015 bp DNA linear HTG 31-AUG-2001  
 LOCUS Homo sapiens chromosome 5 clone CTD-2385B5, WORKING DRAFT SEQUENCE,  
 DEFINITION 4 unordered pieces.  
 AC091870  
 AC091870.2 GI:15290419  
 VERSION HTG: HTGS-PHASE1; HTGS-DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS human.  
 SOURCE Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 99015)

## AUTHORS

DOE Joint Genome Institute.

## TITLE

Sequencing of Human Chromosome 5

## JOURNAL

Unpublished

## REFERENCE

2 (bases 1 to 99015)

## AUTHORS

DOE Joint Genome Institute.

## TITLE

Direct Submission

## JOURNAL

Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint

## COMMENT

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 On Aug 25, 2001 this sequence version replaced gi:14333806.

-----Genome Center

Center: Joint Genome Institute

Center code: JGI

Web site: <http://www.jgi.doe.gov>

-----

Project Information

Center Project Name: 775418

Center clone name: CITB-H1\_2385B5

-----

## Summary Statistics

Consensus quality: 95623 bases at least Q40

Consensus quality: 97794 bases at least Q30

Estimated insert size: 75000; agarose-fp estimation

Estimated insert size: 98715; sum-of-contigs estimation

Quality coverage: 4.94 in Q20 bases; agarose-fp estimation

\* NOTE: this is a 'working draft' sequence. It currently

\* consists of 4 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 11898: contig of 11898 bp in length

\* 11899 11998: gap of unknown length

\* 11999 25054: contig of 13056 bp in length

\* 25055 25154: gap of unknown length

\* 25155 58052: contig of 32898 bp in length

\* 58053 58152: gap of unknown length

\* 58153 99015: contig of 40863 bp in length.

Location/Qualifiers

1. 99015

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="5"

/clone="CTD-2385B5"

/clone\_lib="Caltech human BAC library D"

31380 a 18195 c 18463 g 30676 t 301 others

BASE COUNT 1.2%; Score 20; DB 2; Length 99015;

ORIGIN Best Local Similarity 100.0%; Pred. No. 21;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 584 tcagattgttcaggaggttg 603  
 |||  
 Db 29181 TCAGATTGTTTCAGGAGCTTG 29162

## RESULT 8

AC021694 149179 bp DNA linear PRI 11-DEC-2001  
 LOCUS Homo sapiens, clone RP11-24F3, complete sequence.  
 AC021694  
 AC021694.8 GI:17488694  
 VERSION HTG.  
 KEYWORDS human.  
 SOURCE Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 149179)

## AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

## TITLE

Homo sapiens, clone RP11-24F3

## JOURNAL

Unpublished

## REFERENCE

2 (bases 1 to 149179)

## AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,  
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,  
 Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 DeArelano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J.,  
 Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
 Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Landers, T., Lechoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
 McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,  
 Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 149179)

## REFERENCE

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,  
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,  
 Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A.,  
 Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
 Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,  
 Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
 Meneus, L., Mikova, T., Mianga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,  
 Seaman, S., Severy, P., Sougnéz, C., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,  
 Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (09-AUG-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 149179)

## REFERENCE

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,  
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,  
 Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A.,  
 Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,

Perreira, P., FitzHugh, W., Gage, D., Galagan, J., Cardyna, S., Ginde, S., Gird, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRoque, K., Lamacz, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Rettig, M., Roy, A., Santos, R., Schauer, S., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Rogov, P., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zalnoun, J., Zembek, L., Zimmer, A. and Zody, M.

# TITLE

JOURNAL Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Dec 11, 2001 this sequence version replaced gi:15144324. All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WTHR

Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information

Center project name: L4448  
 Center clone name: 24\_F\_3  
 -----

## FEATURES

### source

1. 149179  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-24F3"  
 /clone\_lib="RP11-11 Human Male BAC"  
 123. 179  
 /rpt\_family="HAL1b"  
 207. 349  
 /rpt\_family="HAL1b"  
 1299. 1603  
 /rpt\_family="AluJo"  
 2392. 2429  
 /rpt\_family="AT\_rich"  
 2551. 2637  
 /rpt\_family="AluY"  
 2669. 2815  
 /rpt\_family="MIR"  
 2880. 3595  
 /rpt\_family="MER50"  
 complement(3603..3730)  
 /rpt\_family="MIR"  
 3753. 4357  
 /rpt\_family="L1R8A"  
 4361. 4487  
 /rpt\_family="MER5B"  
 5030. 5261  
 /rpt\_family="MER39"  
 complement(5262..5675)  
 /rpt\_family="MSTA"  
 5675. 5805  
 /rpt\_family="MER39"  
 6044. 6095  
 /rpt\_family="AT\_rich"  
 7510. 7606  
 /rpt\_family="MER53"  
 7992. 8136  
 /rpt\_family="MER5A"  
 9440. 9629  
 /rpt\_family="MLT1L"

repeat\_region 11620..11663  
 /rpt\_family="(TCA)n"  
 repeat\_region 11763..11845  
 /rpt\_family="MIR"  
 repeat\_region complement(11846..12113)  
 /rpt\_family="AluJB"  
 12114..12231  
 /rpt\_family="MIR"  
 repeat\_region 12300..13976  
 /rpt\_family="MER52A"  
 complement(13977..14057)  
 /rpt\_family="L2"  
 15295..15453  
 /rpt\_family="AluSc"  
 15675..15702  
 /rpt\_family="AT\_rich"  
 15729..15779  
 /rpt\_family="AT\_rich"  
 15925..16013  
 /rpt\_family="AT\_rich"  
 16548..16999  
 /rpt\_family="MER44D"  
 17012..18917  
 /rpt\_family="L1MA3"  
 18937..19287  
 /rpt\_family="THE1B"  
 19288..20649  
 /rpt\_family="THE1B-int"  
 complement(20650..20955)  
 /rpt\_family="AluY"  
 20956..21133  
 /rpt\_family="THE1B-int"  
 21134..21495  
 /rpt\_family="THE1B"  
 21498..26060  
 /rpt\_family="L1MA3"  
 26061..26084  
 /rpt\_family="(TAAAA)n"  
 26085..26236  
 /rpt\_family="L1MA3"  
 26247..32756  
 /rpt\_family="L1PA7"  
 32744..33664  
 /rpt\_family="L1MA3"  
 33676..33925  
 /rpt\_family="MER44D"  
 complement(34724..34815)  
 /rpt\_family="MIR3"  
 complement(35207..35505)  
 /rpt\_family="AluY"  
 complement(35771..35884)  
 /rpt\_family="MIR"  
 35885..36185  
 /rpt\_family="AluY"  
 complement(36186..36262)  
 /rpt\_family="MIR"  
 36361..36388  
 /rpt\_family="AT\_rich"  
 36644..36848  
 /rpt\_family="MER34B"  
 repeat\_region 37019..37063

Query Match 1.2%; Score 20; DB 9; Length 149179;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 584 tccagattgttcaggattg 603  
 |||||

DB 113293 TCCAGATTGTTCCAGAGTTG 113312

RESULT 9  
 AC024978

```

LOCUS       AC024978                      158781 bp    DNA        linear        HTG 26-MAY-2000
DEFINITION  Homo sapiens chromosome 16 clone RP11-285K4 map 16, WORKING DRAFT
SEQUENCE    SEQUENCE, 18 unordered pieces.
ACCESSION   AC024978
VERSION     AC024978.3  GI:8076921
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 158781)
            Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
            Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,
            Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
            Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
            Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
            Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
            Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
            Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
            Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
            Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
            Levine, R., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
            McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
            Meldrum, J., Meneus, L., Mithova, T., Miranda, C., Mienga, V., Morrow, J.,
            Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
            O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierce, N.,
            Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
            Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
            Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
            Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J.,
            Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
            Young, G., Zainoun, J., Zimmer, A. and Zody, M.
  DIRECT SUBMISSION
  Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  On May 25, 2000 this sequence version replaced gi:7249330.
  All repeats were identified using RepeatMasker:
  Smit, A.F.A. & Green, P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html
  ----- Genome Center
  Center: Whitehead Institute/ MIT Center for Genome Research
  Center code: WIBR
  Web site: http://www-seq.wi.mit.edu
  Contact: sequence.submissions@genome.wi.mit.edu
  ----- Project Information
  Center project name: 16961
  Center clone name: 285_K_4
  ----- Summary Statistics
  Sequencing vector: M13; M77815; 100% of reads
  Chemistry: Dye-terminator Big Dye; 100% of reads
  Assembly program: Phrap; version 0.960731
  Consensus quality: 146857 bases at least Q40
  Consensus quality: 151884 bases at least Q30
  Consensus quality: 154688 bases at least Q20
  Insert size: 157081; sum-of-contigs
  Quality coverage: 4.5 in Q20 bases; sum-of-contigs
  -----
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 18 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  *
  * 1 1045: contig of 1045 bp in length
  * 1046 1145: gap of 100 bp
  * 1146 2190: contig of 1045 bp in length
  * 2191 2290: gap of 100 bp

```

```

*
* 2291 3862: contig of 1572 bp in length
* 3863 3962: gap of 100 bp
* 3963 5098: contig of 1136 bp in length
* 5099 5198: gap of 100 bp
* 5199 6666: contig of 1468 bp in length
* 6667 6766: gap of 100 bp
* 6767 9285: contig of 2519 bp in length
* 9286 9385: gap of 100 bp
* 9386 11644: contig of 2259 bp in length
* 11645 11744: gap of 100 bp
* 11745 14644: contig of 2900 bp in length
* 14645 14744: gap of 100 bp
* 14745 16763: contig of 2019 bp in length
* 16764 16963: gap of 100 bp
* 16964 22636: contig of 5773 bp in length
* 22637 22736: gap of 100 bp
* 22737 30382: contig of 7846 bp in length
* 30383 30682: gap of 100 bp
* 30683 36993: contig of 6311 bp in length
* 36994 37093: gap of 100 bp
* 37094 46226: contig of 9133 bp in length
* 46227 46326: gap of 100 bp
* 46327 62414: contig of 16088 bp in length
* 62415 62514: gap of 100 bp
* 62515 81628: contig of 19114 bp in length
* 81629 81728: gap of 100 bp
* 81729 106147: contig of 24419 bp in length
* 106148 106247: gap of 100 bp
* 106248 133039: contig of 26792 bp in length
* 13040 133139: gap of 100 bp
* 133140 158781: contig of 25642 bp in length.
  ----- Location/Qualifiers
  source          1..158781
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="16"
                  /map="16"
                  /clone="RP11-285K4"
                  /clone_lib="RPC1-11 Human Male BAC"
  misc_feature    1..1045
                  /note="assembly_fragment"
  misc_feature    1146..2190
                  /note="assembly_fragment"
  misc_feature    2291..3862
                  /note="assembly_fragment"
  misc_feature    3963..5098
                  /note="assembly_fragment"
  misc_feature    5199..6666
                  /note="assembly_fragment"
  misc_feature    6767..9285
                  /note="assembly_fragment"
  misc_feature    9386..11644
                  /note="assembly_fragment"
  misc_feature    11745..14644
                  /note="assembly_fragment"
  misc_feature    14745..16763
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:right
  misc_feature    16864..22636
                  /note="assembly_fragment"
  misc_feature    22737..30382
                  /note="assembly_fragment"
  misc_feature    30683..36993
                  /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right
  misc_feature    37094..46226
                  /note="assembly_fragment"
  misc_feature    46327..62414
                  /note="assembly_fragment"
  misc_feature    62515..81628
                  /note="assembly_fragment"

```



Mon Apr 29 09:00:24 2002

```

misc_feature      81729..106147
                  /note="assembly_fragment"
misc_feature      106248..133039
                  /note="assembly_fragment"
misc_feature      133140..158781
                  /note="assembly_fragment"
BASE COUNT      45063 a 33044 c 33098 g 45874 t 1702 others
ORIGIN

Query Match      1.2% Score 20; DB 2; Length 158781;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 590 ttgttcaggagttggccaag 609
      |||||
Db 122532 TTGTCAGAGTTGGCCCAAG 122551

RESULT 10
AC067844
LOCUS      AC067844      160099 bp      DNA      linear      PRI 05-NOV-2001
DEFINITION Homo sapiens chromosome 8, clone RP11-234A3, complete sequence.
ACCESSION   AC067844
VERSION     AC067844.6 GI:16041563
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 160099)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,
Boguslavsky,L., Bouckghalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domiano,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczký,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 160099)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckghalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Katat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczký,J., Levine,R., Liu,G.,
McClean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9372
Center clone name: 234_A.3
-----
FEATURES             Location/Qualifiers
     source            1..160099
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
     repeat_region     /clone="RP11-234A3"
                     /clone_lib="RPC1-11 Human Male BAC"
                     /complement(193..790)
                     /rpt_family="LINED"
     repeat_region     836..960
                     /rpt_family="L3"
     repeat_region     1597..1665
                     /rpt_family="(TCTA)n"
     repeat_region     1786..1911
                     /rpt_family="MIR"
     repeat_region     2833..2957
                     /rpt_family="L3"
                     /complement(2966..3033)
     repeat_region     /rpt_family="L2"
     repeat_region     3240..3423
                     /rpt_family="MIR"
     repeat_region     /complement(3419..3647)
                     /rpt_family="MIR"
     repeat_region     3688..3786
                     /rpt_family="L3"
                     /complement(4505..4597)
     repeat_region     /rpt_family="MIR3"
     repeat_region     4849..5029
                     /rpt_family="MIR"
     repeat_region     5907..5937
                     /rpt_family="AT-rich"
     repeat_region     /complement(6084..6265)
                     /rpt_family="MIR"
     repeat_region     6351..6560
                     /rpt_family="L2"
     repeat_region     6678..6858
                     /rpt_family="MIR"
     repeat_region     6912..6961
                     /rpt_family="MIR"
     repeat_region     7039..7140
                     /rpt_family="MIR3"
                     /complement(7896..8087)
     repeat_region     /rpt_family="MIR"
     repeat_region     8336..8536
                     /rpt_family="MIR"
     repeat_region     8545..8685
                     /rpt_family="MIR3"

```

```

repeat_region complement(8854..9028)
/rpt_family="MIR"
repeat_region complement(9975..10085)
/rpt_family="LIR16A1"
repeat_region complement(10455..10875)
/rpt_family="MLTIC"
repeat_region 10921..11295
/rpt_family="HEIC"
repeat_region complement(11667..13061)
/rpt_family="Tigger3b"
13965..14136
/rpt_family="MIR"
repeat_region complement(14340..14437)
/rpt_family="MIR"
15751..15892
/rpt_family="MIR"
repeat_region 15895..16156
/rpt_family="ALUSX"
repeat_region 16256..16307
/rpt_family="(TA)n"
repeat_region complement(16349..16423)
/rpt_family="MIR3"
repeat_region complement(16564..16741)
/rpt_family="MIR3"
17600..17714
/rpt_family="MIR3"
repeat_region 18081..18184
/rpt_family="MIR"
repeat_region 18725..18764
/rpt_family="GA-rich"
repeat_region 18765..18775
/rpt_family="MIR"
repeat_region 18776..18945
/rpt_family="MIR"
repeat_region 18946..19053
/rpt_family="MIR"
19153..19188
/rpt_family="AT-rich"
repeat_region 20354..20471
/rpt_family="MERS8A"
repeat_region 20509..20597
/rpt_family="MERS8A"
repeat_region 20852..20943
/rpt_family="(GAA)n"
21012..21146
/rpt_family="MIR"
repeat_region complement(22308..22476)
/rpt_family="MIR"
repeat_region complement(22680..22879)
/rpt_family="MERS8A"
repeat_region complement(23196..23386)
/rpt_family="MIR"
24110..24133
/rpt_family="AT-rich"
repeat_region complement(24310..24871)
/rpt_family="MLTIF1"
24872..24966
/rpt_family="MER103"
repeat_region complement(24990..25134)
/rpt_family="MERS8A"
repeat_region complement(25857..26038)
/rpt_family="MIR"
repeat_region complement(26048..26172)
/rpt_family="MER94"
repeat_region complement(26816..26970)
/rpt_family="MIR"
27132..27473
/rpt_family="L3"
repeat_region complement(28058..28232)
/rpt_family="MIR"
repeat_region complement(28582..28745)
/rpt_family="MIR"
repeat_region complement(28900..28997)

```

---

```

repeat_region /rpt_family="L2"
29201..29233
Query Match 1.2% Score 20; DB 9; Length 160099;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 842 ccttaattttctgttcg 861
|||||
Db 14486 CCTTAATTTCTGCTGC 144505
|||||

RESULT 11
AC046201/c
LOCUS
DEFINITION 162054 bp DNA linear HTG 12-MAY-2000
HOMO sapiens chromosome 17 clone RP11-199J4 map 17, WORKING DRAFT
SEQUENCE 35 unordered pieces.
AC046201
VERSION AC046201.2 GI:7770643
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 162054)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Landet,E.
TITLE Homo sapiens chromosome 17, clone RP11-199J4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 162054)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bode,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galgani,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehotsky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,J.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Melgrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Navlot,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Traversa,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 12, 2000 this sequence version replaced gi:7547250.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 199_J4
Center clone name: 199_J4
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 140036 bases at least Q40
Consensus quality: 151630 bases at least Q30
Consensus quality: 155993 bases at least Q20
Insert size: 170000; agarose-pp

```

Insert size: 158654; sum-of-contigs  
Quality coverage: 3.1 in Q20 bases; agarose-fp  
Quality coverage: 3.4 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 35 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1391: contig of 1391 bp in length  
\* 1392 1491: gap of 100 bp  
\* 1492 2611: contig of 1120 bp in length  
\* 2612 2711: gap of 100 bp  
\* 2712 3860: contig of 1149 bp in length  
\* 3861 3960: gap of 100 bp  
\* 3961 5294: contig of 1334 bp in length  
\* 5295 5394: gap of 100 bp  
\* 5395 6549: contig of 1155 bp in length  
\* 6550 6649: gap of 100 bp  
\* 6650 8183: contig of 1534 bp in length  
\* 8184 8283: gap of 100 bp  
\* 8284 8813: contig of 530 bp in length  
\* 8814 8913: gap of 100 bp  
\* 8914 10307: contig of 1394 bp in length  
\* 10308 10407: gap of 100 bp  
\* 10408 12790: contig of 2383 bp in length  
\* 12791 12890: gap of 100 bp  
\* 12891 14735: contig of 1845 bp in length  
\* 14736 14835: gap of 100 bp  
\* 14836 16922: contig of 2087 bp in length  
\* 16923 17022: gap of 100 bp  
\* 17023 19071: contig of 2049 bp in length  
\* 19072 19171: gap of 100 bp  
\* 19172 21191: contig of 2020 bp in length  
\* 21192 21291: gap of 100 bp  
\* 21292 24588: contig of 3297 bp in length  
\* 24589 24688: gap of 100 bp  
\* 24689 28907: contig of 4219 bp in length  
\* 28908 29007: gap of 100 bp  
\* 29008 31119: contig of 2112 bp in length  
\* 31120 31219: gap of 100 bp  
\* 31220 34283: contig of 3064 bp in length  
\* 34284 34383: gap of 100 bp  
\* 34384 38785: contig of 4402 bp in length  
\* 38786 38885: gap of 100 bp  
\* 38886 43323: contig of 4438 bp in length  
\* 43324 43423: gap of 100 bp  
\* 43424 46470: contig of 3047 bp in length  
\* 46471 46570: gap of 100 bp  
\* 46571 49103: contig of 2533 bp in length  
\* 49104 49203: gap of 100 bp  
\* 49204 54054: contig of 4851 bp in length  
\* 54055 54154: gap of 100 bp  
\* 54155 57940: contig of 3786 bp in length  
\* 57941 58040: gap of 100 bp  
\* 58041 62511: contig of 4471 bp in length  
\* 62512 62611: gap of 100 bp  
\* 62612 68713: contig of 6102 bp in length  
\* 68714 68813: gap of 100 bp  
\* 68814 74013: contig of 5200 bp in length  
\* 74014 74113: gap of 100 bp  
\* 74114 79000: contig of 4887 bp in length  
\* 79001 79100: gap of 100 bp  
\* 79101 85277: contig of 6177 bp in length  
\* 85278 85377: gap of 100 bp  
\* 85378 92751: contig of 7374 bp in length  
\* 92752 92851: gap of 100 bp  
\* 92852 100166: contig of 7315 bp in length  
\* 100167 100266: gap of 100 bp  
\* 100267 110782: contig of 10516 bp in length

\* 110783 110882: gap of 100 bp  
\* 110883 120476: contig of 9594 bp in length  
\* 120477 120576: gap of 100 bp  
\* 120577 132125: contig of 11549 bp in length  
\* 132126 132225: gap of 100 bp  
\* 132226 145660: contig of 13435 bp in length  
\* 145661 145760: gap of 100 bp  
\* 145761 162054: contig of 16294 bp in length.  
FEATURES  
source  
1. 162054  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="Rp11-199J4"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 1391  
/note="assembly\_fragment"  
1492. 2611  
/note="assembly\_fragment"  
2712. 3860  
/note="assembly\_fragment"  
3961. 5294  
/note="assembly\_fragment"  
5395. 6549  
/note="assembly\_fragment"  
6650. 8183  
/note="assembly\_fragment"  
8284. 8813  
/note="assembly\_fragment"  
clone\_end:17  
vector\_side:right  
8914. 10307  
/note="assembly\_fragment"  
10408. 12790  
/note="assembly\_fragment"  
12891. 14735  
/note="assembly\_fragment"  
14836. 16922  
/note="assembly\_fragment"  
17023. 19071  
/note="assembly\_fragment"  
19172. 21191  
/note="assembly\_fragment"  
21292. 24588  
/note="assembly\_fragment"  
24689. 28907  
/note="assembly\_fragment"  
29008. 31119  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right  
31220. 34283  
/note="assembly\_fragment"  
34384. 38785  
/note="assembly\_fragment"  
38886. 43323  
/note="assembly\_fragment"  
43424. 46470  
/note="assembly\_fragment"  
46571. 49103  
/note="assembly\_fragment"  
49204. 54054  
/note="assembly\_fragment"  
54155. 57940  
/note="assembly\_fragment"  
58041. 62511  
/note="assembly\_fragment"

Query Match 1.2%; Score 20; DB 2; Length 162054;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 590 ttgttcaggagttggccaag 609
DB 65481 TTGTTTCAGGAGTTGGCCAAG 65462

RESULT 12
AC091097 162515 bp DNA linear HTG 09-FEB-2002
LOCUS Homo sapiens chromosome 11 clone RP11-65C9 map 11, *** SEQUENCING
DEFINITION IN PROGRESS ***, 1 ordered pieces.
ACCESSION AC091097
VERSION AC091097.8 GI:118642752
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 162515)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,M., Anderson,S.,
Barne,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,
Camata,J., Campopiano,A., Chang,J., Choepel,J., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Macdonald,P., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheters,R., Meldrim,J., Meneus,C.,
Norman,C.H., O'Connor,T., Murphy,T., Navlor,J., Nguyen,C., Norbu,C.,
Peterson,K.H., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,D.,
Rosetti,M., Roy,A., Santos,R., Schauer,B., Schupback,R., Seaman,S.,
Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye.W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 9, 2002 this sequence version replaced gi:17149471.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11779
Center clone name: 65C_9
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 162515: contig of 162515 bp in length.
* Location/Qualifiers
  1 . 162515
    /organism="Homo sapiens"
    /db_xref="taxon:9606"

FEATURES
source
  /chromosome="11"
  /map="11"
  /clone="RP11-65C9"
  /clone_lib="RPCI-11 Human Male BAC"
  53925 a 30290 c 28951 g 49349 t

Query Match 1-28: Score 20; DB 2: Length 162515;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 584 tccagattgttcaggagttg 603
DB 6630 TCCAGATTGTTTCAGGAGTTG 6649

RESULT 13
AC079184/c 163452 bp DNA linear HTG 15-JAN-2001
LOCUS Homo sapiens chromosome 8 clone RP11-296E5 map 8, WORKING DRAFT
DEFINITION SEQUENCE, 9 unordered pieces.
ACCESSION AC079184
VERSION AC079184.2 GI:12229307
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 163452)
Birren,B., Linton,L., Nusbaum,C., Lander,E.
Homo sapiens chromosome 8, clone rp11-296E5
2 (bases 1 to 163452)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barne,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Burckett,G., Campopiano,A., Cascie,A.,
Choepel,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenqa,V.,
Morrow,J., Murphy,T., Navlor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Oliver,J., Oliver,J., Peterson,K.,
Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Trigglio,J., Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo.A.,
Wilson,B., Wu,X., Wyman,D., Ye.W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-AUG-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 15, 2001 this sequence version replaced gi:9886020.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10551
Center clone name: 296_E5
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

```

Consensus quality: 160037 bases at least Q40  
Consensus quality: 161862 bases at least Q30  
Consensus quality: 162334 bases at least Q20  
Insert size: 163000; agarose-fp  
Quality coverage: 5.7 in Q20 bases; agarose-fp  
Quality coverage: 5.7 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 9659: contig of 9659 bp in length  
\* 9660 9759: gap of 100 bp  
\* 9760 10810: contig of 1051 bp in length  
\* 10811 10910: gap of 100 bp  
\* 10911 14021: contig of 3111 bp in length  
\* 14022 14121: gap of 100 bp  
\* 14122 19568: contig of 5447 bp in length  
\* 19569 19668: gap of 100 bp  
\* 19669 25091: contig of 5423 bp in length  
\* 25092 25191: gap of 100 bp  
\* 25192 31459: contig of 6268 bp in length  
\* 31460 31599: gap of 100 bp  
\* 31560 38963: contig of 7404 bp in length  
\* 38964 39063: gap of 100 bp  
\* 39064 122447: contig of 83384 bp in length  
\* 122448 122547: gap of 100 bp  
\* 122548 163452: contig of 40905 bp in length.

FEATURES

source

1. .163452  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-296E5"  
/clone\_lib="RPC1-11 Human Male BAC"

misc\_feature

1. .9659  
/note="assembly\_fragment  
vector\_side:left"

misc\_feature

9760. .10810  
/note="assembly\_fragment"

misc\_feature

10911. .14021  
/note="assembly\_fragment"

misc\_feature

14122. .19568  
/note="assembly\_fragment"

misc\_feature

19669. .25091  
/note="assembly\_fragment"

misc\_feature

25192. .31459  
/note="assembly\_fragment"

misc\_feature

31560. .38963  
/note="assembly\_fragment"

misc\_feature

39064. .122447  
/note="assembly\_fragment"

misc\_feature

122548. .163452  
/note="assembly\_fragment  
vector\_side:right"

BASE COUNT 52229 a 30041 c 28914 g 51467 t 801 others  
ORIGIN

Query Match 1.2%; Score 20; DB 2; Length 163452;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 584 tccagattgttcaggagttg 603  
|||||

Db 130687 TCCAGATTGTTCCAGAGTTG 130668  
AC051656  
LOCUS RESULT 14  
DEFINITION Homo sapiens chromosome 11 clone RP11-65C9, WORKING DRAFT SEQUENCE,  
164436 bp DNA linear HTG 01-SEP-2000  
24 unordered pieces.  
AC051656  
AC051656 3 GI:9838050  
KEYWORDS HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 164436)  
Waterston,R.H.  
Direct Submission  
Submitted (15-APR-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Aug 17, 2000 this sequence version replaced gi:7637327.

COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NRH065C09  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Sequencing method: plasmid; 0%  
Chemistry: Dye-primer ET; 100% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 150567 bases at least Q40  
Consensus quality: 154488 bases at least Q30  
Consensus quality: 156732 bases at least Q20  
Insert size: 161000; agarose-fp  
Insert size: 163307; sum-of-contigs  
Quality coverage: 3.76 in Q20 bases; agarose-fp  
Quality coverage: 3.78 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 24 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1489: contig of 1489 bp in length  
\* 1490 1589: gap of unknown length  
\* 1590 3247: contig of 1658 bp in length  
\* 3248 3347: gap of unknown length  
\* 3348 5244: contig of 1897 bp in length  
\* 5245 5344: gap of unknown length  
\* 5345 7107: contig of 1763 bp in length  
\* 7108 7207: gap of unknown length  
\* 7208 8923: contig of 1716 bp in length  
\* 8924 9023: gap of unknown length  
\* 9024 10656: contig of 1633 bp in length  
\* 10657 10756: gap of unknown length  
\* 10757 12271: contig of 1515 bp in length  
\* 12272 12371: gap of unknown length  
\* 12372 15389: contig of 3018 bp in length  
\* 15390 15489: gap of unknown length  
\* 15490 19134: contig of 3645 bp in length  
\* 19135 19234: gap of unknown length

```

19235 * 22902: contig of 3668 bp in length
22903 * 23002: gap of unknown length
26247 * 26347: contig of 3245 bp in length
26348 * 26347: gap of unknown length
31074 * 31073: contig of 4726 bp in length
31174 * 31173: gap of unknown length
36939 * 36938: contig of 5765 bp in length
37039 * 37038: gap of unknown length
42664 * 42663: contig of 5625 bp in length
42764 * 42763: gap of unknown length
49841 * 49841: contig of 7078 bp in length
49941 * 49941: gap of unknown length
57986 * 57986: contig of 8045 bp in length
58086 * 58086: gap of unknown length
58087 * 58087: contig of 7408 bp in length
65495 * 65494: gap of unknown length
65595 * 65595: contig of 7599 bp in length
73194 * 73193: gap of unknown length
73294 * 73294: contig of 8087 bp in length
81381 * 81380: gap of unknown length
81481 * 81480: gap of unknown length
89321 * 89321: contig of 7841 bp in length
89322 * 89321: gap of unknown length
89422 * 89422: contig of 11179 bp in length
100601 * 100600: gap of unknown length
100701 * 100700: gap of unknown length
113488 * 113488: contig of 12788 bp in length
113489 * 113488: gap of unknown length
113589 * 113588: contig of 15215 bp in length
128804 * 128803: gap of unknown length
128904 * 128903: contig of 35533 bp in length.
128904 * 128904: contig of 35533 bp in length.
FEATURES
source
1. .164436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-65C9"
1. .1489
/notes="assembly_name:Contig6"
1590. .3247
/notes="assembly_name:Contig7"
3348. .5244
/notes="assembly_name:Contig8"
clone_end:SP6
vector_side:left"
5345. .7107
/notes="assembly_name:Contig9"
7208. .8923
/notes="assembly_name:Contig10"
9024. .10656
/notes="assembly_name:Contig11"
10757. .12271
/notes="assembly_name:Contig12"
12372. .15389
/notes="assembly_name:Contig13"
15490. .19154
/notes="assembly_name:Contig14"
19235. .22902
/notes="assembly_name:Contig15"
23003. .26247
/notes="assembly_name:Contig16"
clone_end:T7
vector_side:left"
26348. .31073
/notes="assembly_name:Contig17"
31174. .36938
/notes="assembly_name:Contig18"
37039. .42663
/notes="assembly_name:Contig19"
42764. .49841
/notes="assembly_name:Contig20"
49942. .57986
/notes="assembly_name:Contig21"
58087. .65494
/notes="assembly_name:Contig22"
65595. .73193
/notes="assembly_name:Contig23"
73294. .81380
/notes="assembly_name:Contig24"
81481. .89321
/notes="assembly_name:Contig25"
89422. .100600
/notes="assembly_name:Contig26"
100701. .113488
/notes="assembly_name:Contig27"
113589. .128803
/notes="assembly_name:Contig28"
128904. .164436
/notes="assembly_name:Contig29"
BASE COUNT 52702 a 30327 c 28951 g 50143 C 2313 others
ORIGIN
Query Match 1.2%; Score 20; DB 2; Length 164436;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 584 tccagattgttcaggagttg 603
|||||
Db 30757 TCCAGATTGTCAGGAGTTG 30776
|||||
RESULT 15
AC087565/c 192974 bp DNA linear PRI 04-FEB-2002
LOCUS Homo sapiens chromosome 16 clone RP11-53L24, complete sequence.
DEFINITION AC087565
ACCESSION AC087565
VERSION AC087565.4 GI:18483429
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 192974)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 192974)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 192974)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Feb 4, 2002 this sequence version replaced gi:15963645.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----

```

us-09-882-694a-2.rge

• Mon Apr 29 09:00:24 2002

FEATURES  
source  
Location/Qualifiers  
1. .192974  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-53L24"  
BASE COUNT 57606 a 41225 c 39814 g 54329 t  
ORIGIN

Query Match 1.2% Score 20; DB 9; Length 192974;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 590 ttgttcaggaggttgccaag 609  
|||||  
Db 71241 TTGTCAGGAGCTTGCCCAAG 71222

Search completed: April 27, 2002, 05:23:55  
Job time: 16672 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 07:57:47 ; Search time 114.27 Seconds  
(without alignments)  
1912.073 Million cell updates/sec

Title: US-09-882-694A-11  
Perfect score: 6455  
Sequence: 1 MADESEKPRNQDSESSSH.....LQRLGRYVEMCLAQSLDOA 1263

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	2378	36.8	1284	3 Q9Y8G2	Q9Y8G2 emericella
2	2001	31.0	1348	3 Q9Y8G1	Q9Y8G1 emericella
3	1971.5	30.5	1331	3 Q9HGT5	Q9HGT5 trichophyto
4	1968.5	30.5	1293	3 Q9Y748	Q9Y748 emericella
5	1956	30.3	1349	3 Q43121	Q43121 aspergillus
6	1789.5	27.7	1408	3 Q43140	Q43140 cryptococcus
7	1780	27.6	1307	3 Q43138	Q43138 aspergillus
8	1715	26.6	1288	13 Q93437	Q93437 gallus gall
9	1656.5	25.7	1287	13 Q91586	Q91586 xenopus lae
10	1643.5	25.5	1280	6 Q46605	Q46605 canis fam
11	1640	25.4	1285	6 Q02793	Q02793 ovis aries
12	1637.5	25.4	1272	11 Q91K64	Q91K64 rattus norv
13	1629.5	25.2	1272	11 Q924L6	Q924L6 rattus norv
14	1601.5	24.8	1294	5 Q45721	Q45721 caenorhabdi
15	1594	24.7	1292	10 Q9MI09	Q9MI09 arabidopsis
16	1585.5	24.6	1278	10 Q9FWX7	Q9FWX7 arabidopsis

Q24850 entamoeba h  
Q9V626 drosophila  
Q941H6 coptic japo  
Q9512 arabidopsis  
Q93149 arabidopsis  
Q919X1 oryza sativ  
Q21349 caenorhabdi  
Q80725 arabidopsis  
Q9735 caenorhabdi  
Q04711 arabidopsis  
Q60502 cricetus  
Q9C163 rhizomucor  
Q20331 caenorhabdi  
Q95Y13 arabidopsis  
Q9C7F8 arabidopsis  
Q9FHF1 arabidopsis  
Q24393 drosophila  
Q9FWX8 arabidopsis  
Q61301 haemochus  
Q90435 raja erinac  
Q92IG2 solanum tub  
Q24851 entamoeba h  
Q01495 caenorhabdi  
Q913X0 arabidopsis  
Q20335 caenorhabdi  
Q95472 caenorhabdi  
Q9736 caenorhabdi  
Q9V616 drosophila  
Q9SDM5 gossypium h

ALIGNMENTS

RESULT 1  
ID Q9Y8G2 PRELIMINARY: PRT; 1284 AA.  
AC Q9Y8G2;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ATP-BINDING CASSETTE MULTIDRUG TRANSPORT PROTEIN ATRC.  
GN ATRC.  
OS Emericella nidulans (Aspergillus nidulans).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiales; Trichocomaceae; Emericella.  
OX NCBI\_TaxID:5072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-W6-096;  
RX MEDLINE=20408170; PubMed=10954082;  
RA Andrade A.C., Van Nistelrooy J.G.W., Perry R.B., Skatrud P.L.,  
RA De Waard M.A.;  
RT "The role of ABC transporters from Aspergillus nidulans in protection  
RT against cytotoxic agents and in antibiotic production.";  
RL Mol. Gen. Genet. 263:966-977(2000).  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
DR EMBL; AF071410; AAA43625.1; -.  
DR HSSP; P13569; INBD.  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR001140; ABC\_transporter\_tmem.  
DR InterPro; IPR003439; ABC\_transporter.  
DR InterPro; IPR001687; ATP\_GTP\_A.  
DR InterPro; IPR000568; ATP\_synth\_A.  
DR InterPro; IPR000847; HTH\_LysR.  
DR Pfam; PF00664; ABC\_membrane; 2.  
DR Pfam; PF00005; ABC\_tran; 2.  
DR SMART; SM00382; AAA; 2.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
DR PROSITE; PS00449; ATPASE\_A; UNKNOWN\_1.  
DR PROSITE; PS00044; HTH\_LYSR\_FAMILY; UNKNOWN\_1.  
KW ATP-binding; Transporter.  
SQ SEQUENCE 1284 AA; 139648 MW; D81BB710B50D2BE3 CRC64;

Query Match 36.8%; Score 2378; DB 3; Length 1284;  
Best Local Similarity 40.1%; Pred. No. 1.1e-152;  
Matches 528; Conservative 233; Mismatches 453; Indels 104; Gaps 16;

QY 7 KPRPNODGSESS-----SHPPPEKTEGSSISDYLIRIPRVADKYDWTNLVIALICAGSGAS 62  
DB 8 KETPSQDSSTTSVPCTEAPLVEGEZEASFGAYKRIFTFAGRTLLIQLVAILNACASGAG 67  
QY 63 LPLMSLIQSPFNKNYNSGDSPEAFKADVDHFLVFLPFGKGVLTIVTAAITIS 122  
DB 68 IALQNLIFQCFVTVYDFNGISITPADFDNAAELALFVYLGILARLVLSYTNLLITYA 127  
QY 123 AIFFTTLRRVPLECTLRQEWHPDKOSNGAIATXVTNGNRIQTGAELKLVFTVOALSM 182  
DB 128 AYRIVRNIRHAYLKAAJSQEVAYYDFGSGSIAAQAQTSNGKLIQAGASDKLGLFOGLAA 187  
QY 183 FFAFVVALASOMKALITMSVIPALFVLTGICIAIDAAQAEARTIRYSRAAVALAEVLS 242  
DB 188 FVTLISLSRLWCXKWLTLICIPVATIGTGTGVVAAVEAGHETRLIQLIHAQANSFAGILA 247  
QY 243 SIRTVAFYAKKMYEKYDVFLOAQHOEGKKSPNNGVLFSTFPCYIAAIALAFNKGFR 302  
DB 248 GVKAVHAFGMKRLVRKDEYLVFAHVKVKKISPLGLLFSAEVYTIIVLGVLAFWQGIH 307  
QY 303 MYQNGEADVGVFTVLSVTLAATSIISMLAPSGSVVORRIFGSELSIIDKPTOLDPL 362  
DB 308 MFCRGBIGTAGDIFTVLLSVVIASINLTLLAPYSIEFSRAASAAQLFRLIDRESEINPY 367  
QY 363 DPSKGPEGLGIEIONLAPFVPSQAQVLRDNLIPAGKTTALVAGSGSGKSTWVG 422  
DB 368 GREGLPERVLGDVELENVFTSPTRGITVDNLSLAVPAGKTVLGVGSGSGKSTIVG 427  
QY 423 LLERWYLPSSGRILLDGLGQHNKWLRSIRLVQOQEPVLPFGTIFONTANGFMDEQ-R 481  
DB 428 LLERWYLPSSGAIRLDGMLSELNGLVGMRLNRNLVQOQEPVLPFGSGVDFDNTRYGLVTPWE 487  
QY 482 DLPRQOMELVOKACKANGSDVFINELPNYETEGERAGALSGGQRIARISISDP 541  
DB 488 NASREOMERVOEAKLAYAHEISELTGYDTLGERGGLSGGQRIARISVVSOP 547  
QY 542 KILLDEATSDLPKAEKVOEALNRVSKDRITLVIAHKLATVKSAGNIASVIGSKVIVQ 601  
DB 548 KVLILLDEATSDLPKAEKVOEALNRVSKDRITLVIAHKLATVKSAGNIASVIGSKVIVQ 607  
QY 602 GTHHELIEFGCHVAALRAADLGADEQOEHEKTLHEKAAAEARPERA-LERTHTTAT-- 658  
DB 608 GTHESLIAROGVYAGLVKIONLAVN-----ASAHDNVNEGEDEVALLEVTETATRY 661  
QY 659 -----SQAGDLKRRKVPVGTGLGYSLKLCILIMFYEQKNLYWCFTLSTITVLICA 707  
DB 662 PTSIRGRMNSIKDRDDYENHK-----HMDMLAALAYLVRCPELKNAYLVILGLGSGC 715  
QY 708 ATPFGQALLFSLTLVTELSHAAOERADFYILMFVVALGNLVGYTIGTNCVLSQVV 767  
DB 716 ANYPGQAILMSRVVEVTLSDGMLDRKGFYASMLIVLAAGCLCYLAVGYATNTAQHL 775  
QY 768 THRYQAMQFQVLDQIELLDIPQISGALTQSLSALTQOEISANFLIYIVWGQHL 827  
DB 776 SHWFRLLIHDMLKODIOFFDRENTGALVSRIDSYPHAILLELNGYNIAL-VVIAVLQV 834  
QY 828 EQCSTTSLWMTGPGGCVWCTSTPAFGWLPQ-----NSSRDEARSRLG----- 871  
DB 835 VTCGILAIASFMSKGLGVV-----VEGGIPPLVGAGWVRIRVDSRLDROTSKRYGSSSI 888  
QY 872 -----KLCRCWACKRSSYRDPDLIFDSRRPCSPYTLGHVEQGLAKIITQSF 919  
DB 889 ASEAVNAIRTVSSIAIETVLRRYT-EELDHAVSSSVKPMATMI----- 932  
QY 920 FQRCFCFHLSSMEFLAIALCAVDNMLQVSTTOLAFISSWAFCLPVQAAAOVLAYS 979  
DB 933 ---LTCIBYWFQALGFWGCRVLSLGETSMYSFFFAVLSVFFAQAQAQLFQWS 987

QY 980 TSFTKARSANVILWRLTLPTIRETEENKKKPYGCGPVDEIDIEFRYQRORSARVLRG 1039  
DB 988 TSITGINATNYIAWLHQLQPTVRETPENHDKGPGSGAPIANDVRFSPPLRDPADILKG 1047  
QY 1040 VSMTEPQGFQVAVYVASCCKSTLIALSRFVDPYDTSGRISFAHENIAEMSPRLRGHMSL 1099  
DB 1048 VMLKTNKQFFAFVGGSGCKRTWIAHLERFYDPTGTSITDASTLTDINPISIRIVAG 1107  
QY 1100 VQOEPFLVQGSVRNVTAL-----EAEISEELCOGRPLPAPRMWLILSSLY 1145  
DB 1108 VQOEPFLVQGSVRNVTAL-----EAEISEELCOGRPLPAPRMWLILSSLY 1167  
QY 1146 OKALKRLAAOGRMGFGSGQRIARIALIRNPKLILLDEATSDLTQSRBLVQAALDEA 1205  
DB 1168 O-GIYTPAGSGSGSLSGGQRIARIALIRNPKLILLDEATSDLTQSRBLVQAALDEA 1226  
QY 1206 ST--SRTTIAVAHRLSTIRNVDFVFAANGRIAETGTTHAEIQLRGRYVEMCLAQSLD 1261  
DB 1227 ARDGRDLTVAVAHRLSTIKDANVICVFEGCKIAENGTHQELIVRGGLYRRMCEAQLD 1284

RESULT 2  
QY8G1  
ID QY8G1 PRELIMINARY; PRT; 1348 AA.  
AC QY8G1  
DT 01-NOV-1999 (TEMPUREL 12, Created)  
DT 01-NOV-1999 (TEMPUREL 12, Last sequence update)  
DT 01-DEC-2001 (TEMPUREL 19, Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN MDR (ABC-TRANSPORTER).  
GN ATRD OR ABCD.  
OS Emericella nidulans (Aspergillus nidulans).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eutrotales; Trichocomaceae; Emericella.  
OX NCBI\_TaxID=5072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W6-095;  
RA Andrade A.;  
RT \*Two novel ABC transporters from Emericella nidulans.\*;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBAJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Nascimento A.M., Terenzi M.F., Goldman M.H.S., Goldman G.H.;  
RT \*Molecular characterization of ABC-transporter encoding genes in Aspergillus nidulans.\*;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBAJ databases.  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
DR EMBL: AF071411; AAD43626.1; -;  
DR EMBL: AF173826; AAF29805.1; -;  
DR HSP; P13569; INED.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_Grp\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR ATP-binding; Transport.  
KW SEQUENCE 1348 AA; 147467 MW; 7B0506AB631D0218 CRC64;

Query Match 31.0%; Score 2001; DB 3; Length 1348;  
Best Local Similarity 37.5%; Pred. No. 4.8e-127;  
Matches 500; Conservative 230; Mismatches 488; Indels 114; Gaps 29;

QY 2 ADESEKPRNODGSESS-----SHPPPE-----KETEGRISDYLRIFRYADKY 44  
DB 54 ADKHERFKSSSSNNAYSVNEVDALIAHLPEDEQVRLKTOLEEKVNIS-PFGLWRYATKM 112  
QY 45 DWTNLVIALICAGSGASLPLMSLIQFSTFKNNYNSGDSPEAKADVDHFLVFWYL 104



```
Db 241 YKLCALICSSTIVAMILVMG-GISRFVVKSGRMTLVSYGEGTVABEVISSIRNATAFCT 299
QY 253 QKKMVEKYDFVLOQAHOEGKKSPNNGVLFSTEFYCYAIAALAFWKFGRMYQNGEVADV 312
Db 300 OEKLAROEVLHKEARKGRLOMMLGIMFGSMWAIMSYNYGLGFWMGSRFLVGE-TDL 358
QY 313 KGVFTVVLVLAATISMLAPSGSVYQRRIFGSELSFIIDKPTOLDPLDPGSGKOPEGC 372
Db 359 SAIVNILLAIWIGFSIGNVAPNTQAFASAIKAKIFSTIDRVSAIDPCSGDEGPTIEN 418
QY 373 LGQIEIQNLAFAYPSRPSAQVLRDNLTPAGKTTALVGASGSKTMVGLLRLERWLPSS 432
Db 419 EGTIEFGIKHYPSRPEVVMYMEDINLVWPKGTTALVPGSGSKTVVGLLRFYNPVS 478
QY 433 GRILLDGLGLOYNVVKLRSRIRLVQEPVLFRTTIFONTANGFMDQQLPREKOM-EL 491
Db 479 GSVLDGRDITLNLRLRQOISLVSEPTLFGTTIFENIRLGLGSPMENESEFOIKER 538
QY 492 VOKACKASNGDVFINELPNGYETEVERAGALSGGQRQRTAIARSIIDPKILLDEATS 551
Db 539 IVSAAKEANAHDFINGLPDGVATDVGQRFLLSGGQKQRTAIARAIIVSDPKILLDEATS 598
QY 552 ALDPKAEKVQBAALNRVSKDRTTLVIAHKLATVKSAGNIAVISOQKIVBQGHHELIEFG 611
Db 599 ALDTKSEGVQAALDAASGRGTTIVIAHRLSTIKSADNIWVVGRIAEQGHDELVDKK 658
QY 612 CHYAALVRAQDLGAD--EQOHEKTLHEK-----NAREAGERP-----ALERT 653
Db 659 GYTLQVLAQAKINERGSEDEAVLEKEIKSRQISVPAKSVNSKYDPEDVEANLGR 718
QY 654 HTTATSOAGDLEKRPVPGTIGYSLKILIMFYEQKNLYWCFLSTITVLICAATFPQ 713
Db 719 DFKKSLSSVILSKRSQEKETEYSLGTLIRFIAGFNKPERLIMLCGFFFAVLGAGQPVQ 778
QY 714 ALLFSLRLTVFELSGH---AAQERADPVILMFFVVALGNLVGFTIGWTCNVISOVWTR 770
Db 779 SVFFAKGITTLSLPSLVGLKREDANFSLMFLMLGLVQLVYQSAQGVIFAICESLIYR 838
QY 771 YQANFQVRLDODIELDIPQISGALTQSLSALPQLQBELISANFLIYVVGQHRLEQC 830
Db 839 AKSKFRAMLKODIAFFDLPENSTGALTSLFTETHLGSVSGTALTGLTILM----- 890
QY 831 STTSLMETGPGCGVCTSTPAFGWLPONSSRDEAKSRKLGKIC-----RKCW 878
Db 891 STTLI-----VALTVLAFGW-----KALVCISTVPVLLCGFYRFW 928
QY 879 ACKRSYRDPDLIFDSRRPCSPVLGHV-----EQLAKIIOSPFQK----- 922
Db 929 ILAQFQTRAKKAYESSASYACEATSSRTVASLUREQVMEIY-----GQLNDQAKSLR 984
QY 923 -----CFGFHLQSNEFLAIALGCIAYDNWL-----QVSTTQNLNFISSWAFCLPV 969
Db 985 SVAKSSLLYAAQSQSPFCLALGE-----WYGGGILGKEYNAFOFFLCISCVIF--GS 1036
QY 970 QAAQYLAYSTSTPKARSAANYILWLRLKPTIR-ETEENKKKGPVGGCPVLEDIEPRY 1028
Db 1037 QSAGIVFSPDMCKAKSAAADFKRLDPVPTIDIESTGKLETVEG-TIEFRDVHFRY 1095
QY 1029 RORDSARVLRGVMTIEPQGVAVYVAGSGCKSTLALSERFYDPTSGRISPAHENIAEM 1088
Db 1096 PTPRQPVRLGRLNLTVRGQYIALVPGSGCKSTTIALVERFYDVLSSGGVYIDGRDISRL 1155
QY 1089 SPRLYRGHMSLVQOEPHTLYQGSVRENVTALEN-ELSEELCOGLRPARPMLWILSLYQK 1147
Db 1156 NVNYSRHLVLSQEPHTLYQGTITDNLVLLGVDRDDPQVFAACKAANIYDFIMSL-PD 1214
QY 1148 ALKRLAARGMFGSGGQRQRTAIARALIRNPKILLDEATSAIDTQSERLVQAALDEAST 1207
Db 1215 GFTVVGSGMLSGGQKQRTAIARALIRDPKVLLDDEATSAIDTQSERLVQAALDEAST 1274
QY 1208 SRTTIAVHRLSTRNVDPVFPVANGRIARTGTHAEOLRLGRFYEMCLAQSLQD 1262
Db 1275 GRTTIAVHRLSTRNVDPVFPVANGRIARTGTHAEOLRLGRFYEMCLAQSLQD 1329
```

## RESULT 4

```
Q9Y748
ID Q9Y748 PRELIMINARY; PRT: 1293 AA.
AC Q9Y748;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ABC TRANSPORTER PROTEIN ATRC.
GN ATRC.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=FGSC A4;
RX MEDLINE=99156773; PubMed=10036328;
RA Angermayr K., Parson W., Stoffler G., Haas H.;
RT "Expression of atrc encoding a novel member of the ATP binding
RT cassette transporter family in Aspergillus nidulans-is sensitive to
RT cycloheximide.";
RL Blochim. Biophys. Acta 1453:304-310(1999).
CC 1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF082072; AAD25925.1;
DR HSSP: P13569; INED.
DR InterPro: IPR0033593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF001687; ATP_GTP-A.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1293 AA; 141984 MW; 8D13AEDFC9E946D3 CRC64;
```

Query Match 30.5%; Score 1968.5; DB 3; Length 1293;

Best Local Similarity 36.1%; Pred. No. 7.3e-125;

Matches 466; Conservative 243; Mismatches 487; Indels 95; Gaps 22;

```
QY 24 EKETEISDYLRIPRYADKYDWTMLNVIATICAIGSGASPLMSIIFSGFTNKFNNYSG 83
Db 46 DPDROSGTKAYPRVWGYSSVQDHVIRVTGLPAAIACGAALPMMLTVFGTLIDNFDWGAG 105
QY 84 DGSPEAFKADVDHFLVFWFLYFVIGKFLVTVYSTAAITISALRTTTLRRVFELECTROEV 143
Db 106 KLSDFEFSHVSONALNFTYLFISIFVLSSTACLRATRCVRLRHRDFIRSLRQDL 165
QY 144 WHPDKOSNGAIXYVTTNGNIGTQGIABKLVFTVQALSMFFSAPVVALASQWKLALITHS 203
Db 166 SYFDNCLPGTVAIVLSNNADLVETGLGKLGIAIEGVAQICAAAFVAFARONKTLIVAA 225
QY 204 VIPAIFLVGTGICATDAQAQEARITRYSRAAVLAEEVLSSIRTVHAFYAKKMYDYDF 263
Db 226 TLPLAMLVVTVVILETRITTKILAIYSKAGGIAEEALASTHIVKAYNAASKLQARYDSY 285
QY 264 LQQAHOEGKKSPNNGVLFSTEFYCYAIAALAFWKFGRMYQNGEVADVCKVFTVLSVT 323
Db 286 LERATQLVGKRPFGIQYGAQFAIMFCAYALANFYGIRLLVKEIESGGYLTIVTSVL 345
QY 324 LAATSIIMLAPSGSVVYQRRIFGSELSFIIDKPTOLDPLDPGSGKOPEGCQGTIONLAF 393
Db 346 IQGSQSLTLGPIGVSATAAQAQELFOVIDRKNIDLSLSDGRTLTGVTGHSIFRNVSF 405
QY 384 AYPSPRSQVLRDNLTPAGKTTALVGASGSKTMVGLLRLERWLPSSGRILLDGLG 443
Db 406 AYPSPRSVRLDDVTDFEAGKTTAIVGSGSGKSTILALVSRFPDVPVSGVLLDGHPIH 465
QY 444 QYNKWLRSRIRLVQOEPVLFRTTIFONTANGFMDQQLPREKOMELVQKACKSNGD 502
```

```
Db 466 ELNRLRGQIGSVQBPVLFSESI FANVCHGFFRTDMDLLPEHERRIRVOEACEAFAH 525
QY 503 VFINELPNGETVEGERAGALSGQQRQRTAIAKSIISDPKILLDDEATSDALDPKAEKVQ 562
Db 526 HFOGLPQYDTPPGAGLSSGGQRTAIAKSIIRNPILLDDEATSDALDPKAEKVQ 585
QY 563 EALNRVSKDRITLVIAHLKATVKSAGNIATVSGKIVGQTHHELIIEFGCHYALVRAQ- 621
Db 586 AALNVNSTRITVIAHRLSTVORADNVIVVLRGRVVEGSHRELLAKKGTGYFLVAAQT 645
QY 622 ----DLGADE--QOEHEKTLHEKAAREAGERPALERTHTTATSQAGDLEKRVVPGTL 674
Db 646 EDGIISTAGDQGTSEADSVLVKKAIVEDYNSAMSSQSKATA-----HEKRPV---- 694
QY 675 GYSLKILIMFYQKNIYCFLLSTITVILICATFPQOALLF--SKLLIVFELSGHAAQ 732
Db 695 ---LUKCLFILLRGLQMLPFCGLIVLYRCQCLPSSGRFQOHLIFQFPPLPAMASE 751
QY 733 --ERADFYLMFVVALGNLVGYFTIGWTCNVISQVYVTHRYQAAMFORVLDODIELDIP 790
Db 752 MLHKNFWGIIYIVLAVSVLICVAGLGFFTVAASFVLSGTVRSRYFAAMLNQDVSFFEE 811
QY 791 EGISGALTSQLSALPTOLOELIS--ANFLIYVVGHRLEQCSSTSLWMETGGCGVWCT 848
Db 812 DQSGAVMTQSLTDPQRIEDLSICLGFILLVVV--NVLASG---ILALAVG-----WRL 861
QY 849 STPA-FCWLPQNSRDEARSKLCKKCKWACKRSYRDPDLIFDSRRPCSPV---- 903
Db 862 ALVAIFGCLPPLFLAGVYVR-LEITCQ-----ERT-----RLYLSARFATEISAIR 910
QY 904 -----LGHVEGLAKI--IQSEWFGRCFGFHLSSOSMEFLATALGFCI 943
Db 911 TVASITLBEKVIQMYDELRSHTSPKFIITLVSAILLGLC-----ESLYIATLGLFWY 964
QY 944 AYDNMLQYSTQLNFISSSWAFCLPQAAQYLAQYLAQYLAQYLAQYLAQYLAQYLAQY 1003
Db 965 GVKLLSQEYVETFEVAVFVGGQAGFLGTVNTAKTAANNIHLILGRPSIN 1024
QY 1004 ETEENKKGVPV-GGCPVDLEDIEFRYRDSARVLGRYSMTIEPCGFVAVVGASGCGKST 1062
Db 1025 ASTGQKQEVOLDSDAIEFDVHFVSYPARTVPLVGLSLFVKVKEGHEGLVGSAGCKT 1084
QY 1063 LIALSERVDPYSGRISEFAHENTAEPSRLYRGHMSLVQOEPYLVQGSVRENTLA---- 1118
Db 1085 VISLERYEAGSGEIFINGIPLHDIDVHSHRARGLVQNTPLVQGSIRDNLIGISIS 1144
QY 1119 ----LEAESELCQGLFARPMMLWSLSYQKALKRLAAQRGQFSGGQQRQRIATARA 1173
Db 1145 HQNEELDSQIEKLTAKYKQANIOTDFTQS--LPEGQOTDPTGRGLALSGGQQRQRIATARA 1202
QY 1174 LIRNPKLLLDDEATSDALDQSERLVQAALDEAS--TSRTTIAVARELSTIRNVDVIFVPA 1231
Db 1203 LIRDPPELLLDDEATSDALDQSERLVQAALDEAS--TSRTTIAVARELSTIRNVDVIFVPA 1262
QY 1232 NGRIRATGTHAELQRLGRYEMCLAGSLDQ 1262
Db 1263 EGVVEEGTHAELMARGGRYVQWLAQGLDR 1293
RESULT 5
O43121 PRELIMINARY; PRT; 1349 AA.
AC O43121
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MULLIDRUG RESISTANCE PROTEIN 1.
GN MDR1.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_Taxid=5085;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=10AE/86/10;
RA Tobin M.B., Peery R.B., Skatrud P.L.;
RT "Genes encoding multiple drug resistance-like proteins in Aspergillus
RT fumigatus and Aspergillus flavus.";
RL Gene 200:11-23(1997).
CC -/- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; U62934; AAB88658.1; -.
DR EMBL; U62933; AAB88657.1; -.
DR HSP; P13569; INBD.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR001140; ABC_transporter_Lmem.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1349 AA; 147723 MW; B90A86F9D62C4165 CRC64;

Query Match 30.3%; Score 1956; DB 3; Length 1349;
Best Local Similarity 35.8%; Pred. No. 5 5e-124;
Matches 491; Conservative 229; Mismatches 478; Indels 172; Gaps 29;

QY 4 ESEKPRPNQDSESSSHPP- EKETEGISD----- 33
Db 35 DNEKPHDHSLSDTLMAPPDGGKKDHKAVDNDLSLFAHQHEKEVLKRLDAPSVKV 94
QY 34 -YLIRIYRYADKYDWTNLVIALICAGSGLPLMSIFGFTNKNFNYSNGDSPAFA 92
Db 95 SFETLVYASRRDILILVSAICAAGAALPLTILFGLSASAFQISLGTMPHYEYH 154
QY 93 DVDFHFLVPLVYFGKFLVTVVSTAATISAIITRTLRVFELECTLRQEVHFKQSG 152
Db 155 KLTKNVLYFVYLGAEEFVTVVSTVGFYTGHELTQKIRENVLEAILRONMAYPDKLAG 214
QY 153 AIATKYTTNGNRIORGAEKLVFTVQALSMFFSAFVVALASOWKALITMSVIPALFVT 212
Db 215 EVTRITADNLIQDAISEKVLITAFATFVAFVAVYVWKALICTSTIVALNVM 274
QY 213 GICIAIDAOEARITRIVSRAVLAEEVLSSIRTVIAFYAQKKWKEVDYVFOQHQRGK 272
Db 275 GGSRFIVYKSKSTESYGAGTVAEEVISIRNATAFGTDKLAKQVETHLAEAKWG 334
QY 273 KSPNNGVLFSTFYFCIYAAIALAFWKGRMYQNGEADVGRVFTVYVSVTLAATSI 332
Db 335 KQVILGMIMGMFGIMESNTGLGFWMGSRFVVGKEV-NVGQVLTVMILIGSPSLGV 393
QY 333 APSGVVYQRRIFGSELSIIDKPTOLDPLDPSGKOPGCGLOIEQLNLAFAYPSPSAQ 392
Db 394 APNGQAFNGVAAAKIYSTIDRRSPDPYSDGKVLDFEGNIEFRNWKHYIPSPPT 453
QY 393 VLKDFNLITPAGKTTALVAGSGSKSTMWGLLEKRWYLPSSGRILLDGLGQYNNKWLRS 452
Db 454 VMEDVSLSPAGKTTALVPGSGSKSTMWGLVERFVLPVGGVLLDGDHDIOTNLRLWRQ 513
QY 453 RIRLVQOEPVLFRTIFONIANG-----FMDEQRDLPREKQOMELVQKACKASGVDFINE 507
Db 514 QISLVQEPVLFSTTIFRNIHGLIGTKFESKDK---KIRELVENARMANAHDFIWA 569
QY 508 LPNGYETVEGERAGALSGQQRQRTAIAKSIISDPKILLDDEATSDALDPKAEKVQ 567
Db 570 LPEGYDNTVQGRFLLSGQQRQRTAIAKSIISDPKILLDDEATSDALDPKAEKVQ 629
QY 568 VSKDRTTLVIAHLKATVKSAGNIATVSGKIVGQTHHELIIEFGCHYALVRAQ----- 621
Db 630 AAEGRITVIAHRLSTVORADNVIVVLRGRVVEGSHRELLAKKGTGYFLVAAQT 689
QY 622 -----DLGADQOEHEKTLHEKA-----AREAAGERPALERTHT-TATSOAGDLEK 666
```

Db 590 EAEALADADMDADDQEGGVTRITKAVSSNSDVADEKARLEBKMTGTQKSVSSAVLS 749  
Qy 667 RKVPVGLYSLKCI-LIMFYEQKNLYWCFLLSTITVILCAANTFPQALLFSRLITVFE 725  
Db 750 KKVPEQFYSYLTWLVKFTGAFNRPELGYMLGLTGFSL-AGGQPTQAPLYAKAISTLS 808  
Qy 726 LSG---HAQAERADFYLMFFVVALCNLVGVFTTIGWTCNVISQVTVHRYQAAAFQVLDQ 782  
Db 809 LPSEMFHKLHDANFSLMEFVVVGTIAQFISLSINGTAFICSERLIRASQAFRSLRQ 868  
Qy 783 DIELLDIPEICISALTSALPTOLQEL-----ISANFLIYVVGHRLE 828  
Db 869 DISFFRENSGALTSFSLTETKMLSGVGVTLGTIMTSTTLGAAMIALAIGKMLAL 928  
Qy 829 QC-STTSLNMEFGPGVCWCTSTPAFGWLPONSRD-----EARS--RLKGL 873  
Db 929 VCISVPILLACG-----FLRFYMLAQFOORSKAYSAGSASYACEATSIRTVASL 979  
Qy 874 CRK--CWACKRRSYRDPDLIFDSRRPGSPVTLGHVQGLAKIIQSFWFGRCFCFHSQS 931  
Db 980 TREQDVMGVHDQLOKQGR-----KSLISVLRSSL-----YASSQA 1016  
Qy 932 MEFLATALGCIADVMDLVQVSTOLN-----FISSSWAFCLPVQAAAQVLAYSTSF 982  
Db 1017 LVFEFCVALGF-----W--YGTLLGHHEYSIFRFFVCFS-EILFGAQSACTVFSAPDM 1067  
Qy 983 TKARSANVILWRLTKPTIR-ETENKKKGPGVGGCPVDLEDIEFRYRQORSARVLRGS 1041  
Db 1068 GKARNAAGFKLFQSKPTIDWSDEGEKLESMEG-EIEFRDVHFYPTTRPEQPVLRGLN 1126  
Qy 1042 MTIEPGQFVAYGASCGCKSTLIALSERFYDPTSGRISFAHENIAEMSPRLYGRHMSLVQ 1101  
Db 1127 LSVKPGQYALVGPSCGCKSTIALLERFYDALAGGVFVDGKDITKLVNVSYSFSLSVS 1186  
Qy 1102 QEPTLQGSVRENVTALDEA-ELSE---LCQGLRPARPMLWLTLSLYQ-----KALK 1150  
Db 1187 QEPTLYOGTIKENILGYDKDQVSEETLIKVKCD-----ANIYFVMSLEPGD 1235  
Qy 1151 RLAAQGMQFSGQORQTAIARILNPKLLLDENTSALOTSERLVQALDEASTST 1210  
Db 1236 TVVSGKGLMSGQKQORVAIRALRDLKPVLLDEATSAIDSESEKVVQVQALDAAARGT 1295  
Qy 1211 TIAVAHRLSTIRVDVVFVANGRIAEFTGTHAELORLGRGYEMCLAQSL 1360  
Db 1296 TIAVAHRLSTIQNAIIVYFDOGKIVESGTHHELIRNKGRYYELVNLQSL 1345

RESULT 6  
O43140 ID O43140 PRELIMINARY; PRT: 1408 AA.  
AC O43140;  
DT 01-JUN-1998 (TREMblrel. 06, Created)  
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN 1.  
GN MOR1.  
OS Cryptococcus neoformans (Filobasidiella neoformans).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
OX NCBI\_TaxID=5207;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M-106;  
RC MEDLINE=98072425; PubMed=9409767;  
RA Thornevell S.J., Peery R.B., Skatrud P.L.;  
RT "Cloning and characterization of Cnem1: a Cryptococcus neoformans  
RT gene encoding a protein related to multidrug resistance proteins.";  
RL Gene 201:21-25(1997).  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
DR EMBL: U62930; AAC49890.1; -;  
DR EMBL: U62929; AAC49889.1; -;  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR002106; AA\_trna\_ligase\_II.

DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
KW ATP-binding; Transporter.  
SQ SEQUENCE 1408 AA; 152143 MW; CE8A34A09A6F69F8 CRC64;  
  
Query Match 27.7%; Score 1789.5; DB 3; Length 1408;  
Best Local Similarity 34.4%; Pred. No. 1.2e-112;  
Matches 478; Conservative 214; Mismatches 493; Indels 203; Gaps 35;  
  
Qy 6 EKPRNODSSESSHP-----PPEKETEGSI---SDYLRIFRYA 41  
Db 92 EKPISN---AFSKSHPHYKKSKFDLKRKKKEEERKNKEKEASVLPVPSPFALPFA 148  
Qy 42 DKYDWTNLVIALCAIGSCASLPLMSIIFGSEPTNKFNNY-----NSGDGSPA----- 89  
Db 149 APLEIATWLVGLVAAGSCQPLMTLIFGRLTTSFTNYAVIANQISOGGLTPETSALQ 208  
Qy 90 -----PRADVDHPLVFWYLFIFGKFLVYVSTAATISAIRTTTLRRVPLECTLRQEW 144  
Db 209 AKDDDLKTOSGHNLYLMAIGMFLATWLYMFINWVTGELNSKRIRERYLAAYLROEIA 268  
Qy 145 HPKQSGAIAVXTVNGNRIGTIAEKLVFTVQALSMTFFSAFYVALASQKALITMSV 204  
Db 269 YFDLGAAGEVATRIQDCHLVQEGTSEKVALYFOYACTFVCGFVLAFVRSPLAGALVSI 328  
Qy 205 IPAIFVTGCIATAAQAERITRIYSAVALAEVLSSIRTVHAFYAQKWKVEKYDVF 264  
Db 329 LPVIMLCGIMTAMAKYCTAALDHIARAGSLAEVIGSIRTVQAFCKEILGDKFADHI 368  
Qy 265 OQAHQEKKKSPNNGVLFSTVEFCIYAAIALAFKPMYQNGEVAQGVFTVLSVTL 324  
Db 389 EQKIVGRKGSIFEGGLSIMFFVIYAYALAFYGGILVSNCO-ADSGIVINFMVSLI 447  
Qy 325 AATSIKMLAPSGVVYQRRIFGSELSITDKPTOLDLPSPGKQPEGLQIGLQIQLAFA 384  
Db 448 GFSMAMLAPELAAVTKARGAAAKLFATIDRVPAIDSAESEGPKDGLRGEISFENVKPH 507  
Qy 385 YPSRPSAQVLRDFNLITIPAGTKTALVAGSGSGKSTVGLERWYLPSSGRLLDGLQ 444  
Db 508 YPSRPSIPILKGTTFEAGKTFALVAGSGSGKSTVSVLSIERFYDPVSGVVKDGRDIRS 567  
Qy 445 YNVKWLRSRILVVOEPVLFRCITFQNTJANGFMDEQ-RDLPREKQMLVQKACKASNDY 503  
Db 568 LNLNWLROQIGLVQSEPTLFGTTVRGNYVEHGLIGSRYENASLEEKFEVKKACVDANAHN 627  
Qy 504 FINELPNGYETVEGRAGALSGGQRIARISIIIDPKILLIDENTSALDPKAEKVQE 563  
Db 628 FIMKPGQYDTWVGERMILLSGQKQORVAIRAIIVSDPRILLDEATSAJDTSEGVQD 687  
Qy 564 ALNRVSKDRITLIVIAHKLATVKSAGNIAVISQGIKEQGTGHELI--EFGCHYAALVRAQ 621  
Db 688 ALDKASGRRTITIAHRLSTIRADRIYVGGGVEVLQSGSHNDLLANENG-PYAQLVNNQ 746  
Qy 622 DL-----CADRQOEHEKTLHE-----RA 639  
Db 747 KLAQAAAAENLVQDDDIEDPDDAVFVGSSPMQEKQLHRAVTVGRSLASIAMDDIOAKR 806  
Qy 640 AREAAGERPALERTHTTATSQAGDLEKPKVPVGLTGLYLLKCLIMFYEQKNLYWCFLLS 699  
Db 807 AEEVAGE-----DKIP---SSGLYARLLRMSADKPIY---IIA 840  
Qy 700 TITVLCATATFPGOALLFSRLITVVELSG-----HAAQERADFYLMFFVVALG----- 748  
Db 841 FIAACAGWYFSLAILFGKALSDFEIQDPAELRHLSRSA----LWYFTALAAAFVIF 896  
Qy 749 -NLVGYFTIGWTCNVISQVTVHRYQAAAFQVLDQIELLDIPQI1SGALTSQLSALPTQ 807

```
Db 897 FOSAGFSRAGWDLNGV-----LRKKLTATLRHDIEWFDEBERNSTGAVTSLNADQPK 949
Qy 808 LOELISANFLIYVGGHRLQECSTSTSLWMETPGGC-----VWCTSTPAF-- 853
Db 950 VOGFGPTLTGVV-----QSCATLI-----GGCIIGLCYGPLALIGIAPILVS 995
Qy 854 -GWT-----PONSROBARSRKGLKCRKWCACKSSRDPRDLIFDSRRPCSTVL 904
Db 996 GGYIRLVKVVYLKQRMKKLHAASAHLASAAGAVTASLTREX-----DVRRIYSBALK 1050
Qy 905 GHVQGLAKTIQSFWRGCRFGHLSQSMFLAIALGFCIAYDNWL--QVSTTOLNFISS 961
Db 1051 APMKLNFTSIKS-----QCL-FAASOGLTFCIALVFFVIGA-LWIIDAKYSTASTFYVLN 1104
Qy 962 SWAFCLPVOAAOYLAYSTSTFKARSAANYILWLRTLKPTLR-EPEENK---KKGPPVGGC 1017
Db 1105 SIVFA-SIQ-AGNVFFVPDASKANSSAASIFRSIDNEPAINAESNEGVLDHKKVVG-- 1160
Qy 1018 PVDLEDIEFRYRORDSARVLRGVSWTIEPGQPVAVVAGSGCKSTLIALSERFYDPTSGR 1077
Db 1161 HVRIEGVHFRYRPTPGVRLRNUTIDVPAGTVVALVGPSCGCKSTIQMLERFYDPLAGR 1220
Qy 1078 ISPAHENIAEMSRPLRYGHMSLVQOQPTLYOGSVRENTVLAELARLSBELCQGLPAR-- 1135
Db 1221 VTLDGIDIKELNLASYRSQISLVSQEPTLYAGTIRFNLLGANKPI-BEVTODEIDAACK 1279
Qy 1136 -PMLWILLSSYLQALKLAAQCMQFSGGQRORIALARALRNPKLLLLDEATSLDQTS 1194
Db 1280 DANIYDFIVSLPDGFTGVGKGSQSGGQKORIAIARALRNPKVLLDEATSLDQTS 1339
Qy 1195 ERLVOAALDEASTRTTIAVHRLSRINVDVVFVANGRIETGTHAELQRLRGRYEM 1254
Db 1340 EKVYQEAALDKAAKGRTTIATIAHRLSSIOHSRIYFSEGRVAPETHOELLAKKGGYEL 1399
Qy 1255 CLAQLDQ 1262
Db 1400 VOMQNLRS 1407

RESULT 7
ID O43138 PRELIMINARY; PRT: 1307 AA.
AC O43138
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 1.
GN MDR1.
OS Aspergillus flavus.
OC Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_Taxid=5059;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 24133;
RX MEDLINE=98038972; PubMed=9373135;
RA Tobin M.B., Peery R.B., Skatrud P.L.;
RT "Genes encoding multiple drug resistance-like proteins in Aspergillus
RT fumigatus and Aspergillus flavus.";
RL Gene 200;11-23(1997).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: U62932; AAB88656.1; -;
DR EMBL: U62931; AAB88655.1; -;
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmern.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00664; ABC_membrane.2.
DR Pfam: PF00605; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
```

```
SQ SEQUENCE 1307 AA; 143874 MW; 2P65C12CF27FE5F3 CRC64;
Query Match 27.6%; Score 1780; Db 3; Length 1307;
Best Local Similarity 34.3%; Pred. No. 4 8e-112;
Matches 466; Conservative 222; Mismatches 500; Indels 172; Gaps 32;
Qy 7 KPRPNQDSSESS-----SHPPP-----EKTEGSIISDYLRIPRYADKYDWTNLNLTALCA 56
Db 10 KPLPKSPGTGTTTGHVSVAEEVLDQLHTPVQSIGFFGIYRIATRMVDVALFQSAALAA 69
Qy 57 IGSASPLMSIIRGFTNKNFNNSGDGSPFAKADVDHVEFLMFVYLFIFGRFVLYTVYST 116
Db 70 TAGGAALPLFVLFGRLTSTFQDIATHRIYDFHFFHETKNNVYFYIYLGAAEFVAYLAT 129
Qy 117 AAITSAIRTRTRTVFLFLECTLRQEVWHPDKOSNGAIAATXVTNGNRIOTIOTIAEKLVFT 176
Db 130 VGFYTGDDHVVOQIRVEYFOALLRQNLAFDFTLGAAGEITRITVADTNLLODGISEKVGLA 189
Qy 177 VOALSMFSAFVALASQWKLALITMSVIPAIFLVTVGICIAIDAAQAEARITRIYSAVL 236
Db 190 LTGLSTFTVTAIIAYIKMKLALICISALLALLTGWGGCSTLMLFSPKALYQOGRGASM 249
Qy 237 ABEVLSSTRTYHAFYAKKMYKYDYVFLQQAHOEGKKKSPNNGVLFSTFYFCIYAAIALA 296
Db 250 AEDILDSIRTVAAFNAOETLARKYESHKDAEGPMKSKVIFAIMVYGALLCTIYMLNYGLG 309
Qy 297 FMKGFMYQNGEADV--GKVFVTVLSVTLAATSISMLAPSGSVVYVORRIFGSELFISID 354
Db 310 FWMGSRFVEG-TSNIKAGDVLITMMAILGSLYNGVAPNGOALSADAVAANKSLYGTID 368
Qy 355 RPTQLDLPDPSGQPECLGOIEIONLAFAYPSRPSAOVLRDENLTIIPAKKTITLVYVAGS 414
Db 369 RQSPLDALSDQCKTLEFVRGNIVLONIRHYVPSRPEVTVAHDLSCVIPAGKTATVFGPSG 428
Qy 415 SKSTMVGLLERWYLPSSGRITLDGLGQYNVKWLRSIRLWQOEPVIFRGITIFQNTIAN 474
Db 429 SKSTIISLLERFYDPVAGTITMDGHDITQTLNLRWLQOQMSLVSQEPRLFATTAENIRY 488
Qy 475 GFMDQORDLPREKOWEL---VOKACKASNGDVFINELPNGYETEVEGERAGALSGGQRORI 531
Db 489 GLIGSR--FEKSTYEIRKRVAAARMANAHDFIMALPNGYDINI--ESFSLSGGQKORI 544
Qy 532 ATARSIIIDPKILLDEATSAIDPKAEKVQOEALNRVSKORTTLVTAHKLAVKSGNIA 591
Db 545 ATARAIIKDKPILLDEATSAIDPKAEKVQOALDQKASKGRTTIVTAHRLSTIQKAYNII 604
Qy 592 VISQKIVQEQGTHHELIEFGCHYAALVRAQDLGADEQEQEHEKTLHEKAAAREAAGERPALE 651
Db 605 VLANGOIVEQGPHEHLMDBRGIIYCDMVAEHEI-----KKRYSKYSKYSQLLTNLS 655
Qy 652 RHHTTAT-----SQAGDL-----EKKRVPVGTIGYS-LLKILIM 685
Db 656 PKHNPMTFFDKDYPGDDESIIYILSDASDGLHTGEKOR-PVSRMSLSHLMQPVKEE 714
Qy 686 FYEQKNLY-----WCFLLSTITVLICAATP-PGQALLFSRLTYVE---LSGHA 731
Db 715 AYSEWTLKFLASFNRRPEWPELGLGCASLAGIQPSQAVLFKAVASTLSLPLEYKPL 774
Qy 732 QERADFYILMTFFVWALGNLVGYFTIGWTVNVSQVTHRYQAAAFQVRLDQDIELDPIE 791
Db 775 RHDAFWCLMFLMIGIYSLVLYSVQGTLFAYSSKEMVYRARSQAQFRVILHQDISFFDQOE 834
Qy 792 QISGALTSQLSALPTLOE-----LISANFL-----IYIVVGH-----RL 827
Db 835 NTTGALTATLGAAGTKYELGISGVTLGILVSNVNLVSLGVALVIGWKLALVCISNVPAL 894
Qy 828 EOCSTTSLWMTGPGGCVWCTSTPAFGWLQPNSSRDARSRLKGLKLCRCWACK-RSSVR 886
Db 895 LMCGEFVRVWM-----LRFORAKKAYOBSASS-ACEAASAIR 931
Qy 887 DPDFRLIFDS-----RRPCSPTVLGHVEOGLAKIIQSFWFGRCFGFHLSSQSMFLA 936
```

```
Db 932 TVVSLTMEATELQSYQAOLRQLKSDILPIVKSLL-----YASSQALPFC 978
Qy 937 IALGFCIAVDNWL-----QVSTTQNLNFISSWAFCLPVQAAOYLAYSTFYSKARSA 988
Db 979 MALGF-----WYGSLLGHGEYSLFQYFCFSEVIP--GAQAAGTVFVSHAPDMGAKAHA 1030
Qy 989 ANYILWLRLTKPTIRE--TEENKKKG-PVGGCP--VDLEDIEPRYRQRDSARVLRCVSM 1043
Db 1031 A-----REFKRLFSDDTHASRSKGVPTSMRGLVEFRDVSFPRPSRLEQPIRLHLNLT 1084
Qy 1044 IEPGQFVAVYGASGGCTFLALSERFVDTSGTRISFAHENIAEMSPRLYRGHMSLVOQE 1103
Db 1085 IRPGQFVALVAGSGCKSTTALLERYDPLKGVYVYDGNKIITLWMSYRSHLALISGE 1144
Qy 1104 PLYQGSVRENVTALAE--LSEELCQGRLPARPMUWLSLYQAKRLRLAAQGMOPFS 1161
Db 1145 PTLFOGTIRENLTLLGNTPHTVDDFLVKACKDANIYDFILS--LPOGFNTIYVGNKGMLS 1202
Qy 1162 GGORQIAIARALIRNPRLKLLDEATSDLTQSERLVQAAALDEASTSRTTIAVAHRLSTI 1221
Db 1203 GGGQRIARALIRNPRLKLLDEATSDLTQSERLVQAAALDEASTSRTTIAVAHRLSTI 1262
Qy 1222 RNVDFVTFANGRIAECTHAELORLGRYEMCLAQSLD 1261
Db 1263 QRADLIYLDQGEVVESTHRELLRKKGRYELVHLQNDP 1302

RESULT 8
ID O93437 PRELIMINARY: PRT: 1288 AA.
AC O93437;
DC 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ABC TRANSPORTER PROTEIN.
GN CM081.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=99209805; PubMed=10195430;
RA Edelman H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C.,
RA Kane S.E., Kuchler K.;
RT "Cm081", a chicken p-glycoprotein, confers multidrug resistance and
RT interacts with Estradiol.
RL Biol. Chem. 380:231-241(1999).
DR EMBL; AJ009799; CAA08835.1; -.
DR HSP; P13569; INHD.
FT CHAIN 2 1288 ABC TRANSPORTER PROTEIN.
SQ SEQUENCE 1288 AA; 141917 MW; CB258A5F2826DB6C CRC64;

Query Match 26.6%; Score 1715; DB 13; Length 1288;
Best Local Similarity 32.6%; Pred. No. 1.2e-107;
Matches 434; Conservative 243; Mismatches 534; Indels 120; Gaps 22;

Qy 1 MADESEKPRPNQD-----SSSSHPPEKETEG-----SISDYLRIFRYADKYDWT 47
Db 1 MASEDEKQHVVDGTYEIANSSQDDQDEKKGKKGKPPQWSPALFRYSCTDKL 60
Qy 48 LNVIALICAGSGLPLMSIFGSFTNKF-----NNYNSGDSPEAF---KADVD 95
Db 61 LMFISGLLAIAHGTSLPIAMIFGDMTDSFVTGNTNITGSSGLNSSADVFNKLEEMT 120
Qy 96 HVLVFWFLVIGKFLTVVSTAATITSAITRTLRVLECFVROEVVHFQKNGAIA 155
Db 121 RVAYVYSAIAAVLVAAYIQSFVTLAAGRQVKIKREKFFHAIMROEIGNFVNDAGELN 180
Qy 156 TVXTTNGNRIQTGAELVFTVQALSMFFSAFVVALASQWKALITNSVIPALFVLTGIC 215
```

```
Db 181 TRLIDVSNINGIGDKTGLFIQSETTFLTGFIVGFTGKMLTLVILAVSPVGLSALW 240
Qy 216 IADAAQAEARITRIYSRAAVLAEEVLSIRVHYAFYAKRMVKEYDYVFLQAAHQEKKKS 275
Db 241 AKTLTFTDKEAAYAKAGAAVEVLSAVRTVIAFGQOEKEIKRYHKNLDEAKRIGIRKA 300
Qy 276 PNGNVLFSTFEYFCIAAATALAFKGFMYQNGEVADYGVKFTVVLSTLAATSISMLAPS 335
Db 301 ITSINISMGAAFLIYASVALAFWYGTTLILANEYS-IGNVLTVFVSVLIGAFSIGOTAPS 359
Qy 336 GSVVYORRIFGSELESIIDKPTQDLPDPSKOPEGCLGOIEIONLAFAYPSRPAQVLR 395
Db 360 IEAFANARAAYATFNIDNEPEIDSYSDAGHKPHDKGNLEFQNVFFNTPSPRDEVIK 419
Qy 396 DFLNLTIPACKTTALVAGSGCKSTMVGLLERWYLPSSGRILLDGLGLEQYVNVKLSRIR 455
Db 420 GLMLKYNCCQTVVALVGGSCGCKSTTVQIQRFYDPKGETTIDGQDLKSLNVLYLREIIG 479
Qy 456 LVQDEPVLPRGTIFONIANGFMDQORLPREKQMKELVQKACKASNGDVFINELPNGYETE 515
Db 480 VYNOEPVLFATTIAENIRYGRD-----VTNEEIERATKEANAYDFIMKLPKPFETV 531
Qy 516 VGERAGALSGGORQIATARSIIIDPKILLDEATSDLPKAKKYVOEALNRYSKDRTTL 575
Db 532 VGERGAGSGGQRIATARALVHNPKILLDEATSDLTQSERLVQAAALDEAKREGRTTV 591
Qy 576 VTAHLKATVKSAGTAVISQGIQVQTHHELIEFGCHYAAALVRAQDLGADE-----627
Db 592 VVAHRLSTVRNADLIANVFPESGVITBOGNHSQLIEBKGIYKLVNMQTIETEDSSEKSEN 651
Qy 628 -----QOEHEKTLHEKAAREAGERPALERHTHTATSQAGDLKRVKVPVGTLCYSL 678
Db 652 AVSVKESGSSQNLDES-L-KKELRGSSTRSMKKPGEPNDDXGSSPDEELP----PVSF 706
Qy 679 LKCILIMFYEOKNLYWCFLSTITVLICATFPQOALLFSRLITVP-ELSCHAAQGRADF 737
Db 707 LKMLK-----NKNEMYPVAGTFCVAVNGALOPAFSVIFSEIIGIFSETQOKVLRKSNL 762
Qy 738 YILMFFVVALGNLVYFTIGWTCNVISQVTHRYQAAAMFORVLDQIELLDIPEQISGAL 797
Db 763 YSLLEFALGIISFTFFVQGFAGKAGEILTMKLRFAFKMLRODMAMFDDPKNSTGAL 822
Qy 798 TSQLSALPTQLQ-----ELISAN-----FLIYVVGQHRLEQCSSTTSLANETGPG 843
Db 823 TTRLANDASQVKGATGVRLALIAQNIANLGTGIISLVYG-----WQLTLLLVAVP-- 874
Qy 844 CYMCTSTPAFGWLPONSSRDEARSKLGLKCRKC-----WACKRSYRDPDRLLF 893
Db 875 ITAVAGMIEMKMLAGHAKKKELEAAGKIATEAENIRTVASLTREKRFELMYGELLV 934
Qy 894 DSRPCSPVTLGHVBOGLAKIIQSPWFGRCFGFHLSSQSMFLATA-----LGFCIADVNNL 949
Db 935 PYRN-----SYKKAHIFGCFCS--LSQAMFFTYAGCFREGAYLVNNGHI 977
Qy 950 QVSTTQNLNFISSWAFCLPVQAAOYLAYSTFYSKARSAANYILWLRLTKPTIREENK 1009
Db 978 EYKTVFLVPS-----AVVFGAMALGQTSFAPDVAKAKISAAHLFVLPNRPVPPIDSYREDG 1033
Qy 1010 KKGPGVCGPVDLEDIEFRYRORDSARVLRGVSMTEPCOFVAYVYAGSCGCKSTLIALSER 1069
Db 1034 EXPEFGGTRIKDVKNFNPVPEVILQUNLAVKGETALVYVGGSCGCKSTVQVLLER 1093
Qy 1070 FYDPTSGRISFAHENIAEMSPRLYRGHMSLVQOEPTLYQGSVRENVTALAE--ELSEELC 1128
Db 1094 FYDPLSGEIVFDIDAKTLNIOWLRSIHIGIVSQPEILFDTIAENIAYGONSREVSHEEI 1153
Qy 1129 QGRLPARPMUWLSLYQAKRLRLAAQGMOPFSQGORQIATARALIRNPRLKLLDEATSD 1188
Db 1154 ISAAKAASTHSFIDSLPEKYNTVR-GDKGTQLSGGQRIARALIRKPKQLLDEATSD 1212
Qy 1189 ALDTQSERLVQAAALDEASTSRTTIAVAHRLSTIRNVDFVTFANGRIAECTHAELORL 1248
```



Db 1213 ALDTESEKIVQEAALDKAREGRTCTIVIAHRLSTIONADKIAVIONGKIVIEQGTHTQOALLAEK 1272  
QY 1249 GRYEMCLAQS 1259  
Db 1273 GFTYSLVNVQS 1283

RESULT 9  
QY 091586 PRELIMINARY; PRT; 1287 AA.  
AC 091586;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN.  
GN XEMDR.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95322451; PubMed=7599185;  
RA Castillo G., Shen H.J., Horwitz S.B.;  
RT "A homolog of the mammalian multidrug resistance gene (mdr) is  
functionally expressed in the intestine of *Xenopus laevis*.";  
RL Biochim. Biophys. Acta 1262:113-123(1995).  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
DR EMBL: U17608; AAA75000.1; --  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
KW ATP-binding; Transport.  
SQ SEQUENCE 1287 AA; 141505 MW; 06E95073C5771415 CRC64;

Query Match 25.7%; Score 1656.5; DB 13; Length 1287;  
Best Local Similarity 32.4%; Pred. No. 1.2e-103;  
Matches 430; Conservative 231; Mismatches 522; Indels 145; Gaps 24;

QY 10 PNOGSE-----SSSHPPPEKETEGSISDYLRIFRYADKYDWTNLVIALICAIIGSGASLP 64  
Db 22 PMSKKEKGFPSKFKKKEKTEKPKVGVTFMFRYSSTSDKMLMFGTIASLANGAALP 81  
QY 65 LMSIFGFFNKNFNNSGD-----GSPFAKADVDHFLNFFVLFIFGKFLRYV 114  
Db 82 LMLVFGEMTDSFNVGVQDVTGNFTWESMINARELOQMTYAYTSGLGFGVLMCAIY 141  
QY 115 STAATISATHTTTLRRVLECTLOEVWHFDKOSGAIATVTTNGNRIGTGAELV 174  
Db 142 QISFWTLISAGQIKKINSFRAVLRQETGMDINDAGELNRLTDDVSKINEGIGDKIA 201  
QY 175 FTVOALSFFSAFVALASQWIALITMSVIPALFVLTGTCIAIDAAQPARITRIYSRAA 234  
Db 202 MLLQSLTTLVTFGTIGIKGKLTVMGAISPIMGLSAAIWAKVLSAFTNKLKAYAKAG 261  
QY 235 VLAEEVLSIRTVHAFYAKKVKRYDVLQAOHQEGKKSPNNGVLFSTFVFCYAAIA 294  
Db 262 AVAEVLSIRTVAFGQGNKEHRYEKNLDAKKIGIKKAITANVSIGFAFLMIYAAYS 321  
QY 295 LAFNKGPMYONGEVADGVGVTVLVLTATSIISMLAPSGSVVYORIFGSELPSTID 354  
Db 322 LAFWYGTLLIIDGYT-IGSVLTVFAVLIICAFAGVQTSPIEAFANARCAAYTFINIID 380  
QY 355 KPTQDLPDPSGKOPESGICQIPIQNLAFAPYSPRSQVLDDNLTIPACKTTALVGASG 414  
Db 381 NQPKTDSFSGKGLPKDKIGDIEFNKNIPTFYSRKDIQVLKGLNLTIPSGKTVALLVGSSG 440

QY 415 SGKSTWGLLERWYLPSSGRILLDGLGQYNNKWLRSRIRLVQOQEPVLFRTITQIONH 474  
Db 441 CGKSTTVQIQRFYDPEDGVITLDGQDIRSLNIRYUREIIGVVSQEPILPDTTADIRY 500  
QY 475 GEMDEQDLPREKQELVQKAKASNGDVFINELPNQYETEVEGERAGALSGQORIAIA 534  
Db 501 G-----REDVTKEE---IERATKEANAYDFIMKLPDKLETLVGERGTQLSGQORIAIA 552  
QY 535 RSIISDPKILLDEATSSALDPKAEKVVOEALNRVSRDRTTLVIAHKLATVKSAGNIAVIS 594  
Db 553 RALVRNPKILLDEATSSALDTESEAVVQSALDKAREGRTTIIVAHRLSTIRNANAIGFD 612  
QY 595 QGIVSGOQTHLIEFCCHYAALVRAQDLGADQOE-----HEKTL-----HEKAARE 642  
Db 613 NGVIEQSGSHKLMERGCVFNLTQIVTSSKDTEDLETHYIEKKIPVTHTHSNLVR 672  
QY 643 AAGERPALERTHTATSOAGLERKRVPTGLYSLKCLILIMFYEQRLNLCWELLSTIT 702  
Db 673 KSRNTIKSKVPETEDKEVEDEERKK-EEGPPVPVSFFKVMKL-----NPEPYPVVGVC 727  
QY 703 VLICATFFCOALLFSLRTYFELSGHAAQERAD--FYILMFFVVALGNLVGYFTICWT 760  
Db 728 AMINGATQPAFAIIFSRIGVF--AGPVSMRSESSNWSLLFLALGVSFTIFFLQGF 785  
QY 761 NVISQVTHRYOAMFORVLDQDIELLDIPEQISGALTSQLSALPTQLQ-----ELIS 813  
Db 786 GKAGEILTMRUGSFKSMRQEGWFDOSKNTSGALTTLATDASQVOGATGTRALLA 845  
QY 814 AN-----FLIYVVGORHLEQCTTSLAMETPGGCVWCTSTPAFGWLQPNSSRDEAR 866  
Db 846 QNVANLGTALIIISFYG-----WQLTLLILAIVP-VIAAGLVEMKMEFAGHAKKDKKE 897  
QY 867 SRKLGKLC-----RKWACKRSYRQDORLFDSSRRPCSPVLGHVQEQ 910  
Db 898 LERAGKISTDAVINRTVVSLTRKKFERMEKSLSPYR-----937  
QY 911 LAKIOSFWGRCFGPHLSQSMFLAI-----ALGFCIAVDNMLQVSTQLNFISSSWAF 965  
Db 938 -NSIKAHILHGLTYG--LSQAHVLCCLCWFSVLGAYLVVVEGLMKLDEV---FLVSS-AI 990  
QY 966 CLPVOAAQYLAISTSKARSAANYILMLRTLKPTIRETEENKKKGPVGGCPVDLEDIE 1025  
Db 991 VLGAALGOTSSFAFDYTKAMISAAHIFSLRLERVPQIDSYSDQGEKPKCSGNVVFQVN 1050  
QY 1026 FRYRORDSARVLRGYSMTTEPCQFVAVVYGSGGCKSTLIALSERVDPDTSGRISFAHNI 1085  
Db 1051 ENYPTRPDITVLQGLDISVKOGETLALVSGGCKSTTVSILLERFYDPPEGEVLVDGLSV 1110  
QY 1086 AEMSPRLYRGHMSLYOQEPFLVQGSVRENV-----TLAEAELESEELCOG 1130  
Db 1111 RNLNIQWVRAQMGIVSQEPFLFDCSIGDNIAYGDNRRKVTQEEIETAAKEANIHS----- 1165  
QY 1131 RLPARPMILTLSSLYOKALKRLAARGOMPSGGORORAIATARNKLLILLDEATSSAL 1190  
Db 1166 -----FISLTDKYNTRV-GDKTQLSGGOKIAIARALKPKILLDEATSSAL 1215  
QY 1191 DTOSERLVOALDEASTSTTIAVHRLSTIRNVVDVTFPANGRIATGCTHAEQLRGR 1250  
Db 1216 DTSEKVPQALDKMRGTCIVIAHRLSTIONADKIAVIONGKIVIEQGTHTQOALLAEK 1275  
QY 1251 YYEMCLAQ 1258  
Db 1276 YFSLVITIQ 1283

RESULT 10  
QY 046605 PRELIMINARY; PRT; 1280 AA.  
AC 046605;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```
DE MULTIDRUG RESISTANCE P-GLYCOPROTEIN.
GN MDRL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Puel O., Lepage J.F., Alvinerie M., Galtier P., Pineau T.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF045016; AAC02113.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmern.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW ATP-binding; Transport.
SQ SEQUENCE 1280 AA; 141524 MW; 762DD5AFF4C73306 CRC64;

Query Match 25.5%; Score 1643.5; DB 6; Length 1280;
Best Local Similarity 32.2%; Pred. No. 8.7e-103; Indels 141; Gaps 26;
Matches 423; Conservative 246; Mismatches 504;

QY 24 EKETESIDYLRIPRYADKDWLTANLYALICAGSGLPLMSIFGSGFTNKNVNSG 83
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 26 KKEKKPVSTF-AMFYRNNLDRMLVGVGTMAAIIHGAALPLMLLVFGMMTDSFANAGIS 84
QY 84 DGSP-----EAPKADVDHFV-----LMFYLF---IGKFLVTVYSTAATISAIR 126
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 RNTFTPVINESITNTQRFNHLEEMTYAYTYSIGAGVLVAAYIQVSWFVLAAGRQ 144
QY 127 TRLRVFLECLRDVEVWHFQKNGAIAATVTTNGRNIQTGIAEKLFTVQALSMFSA 186
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 145 ILKIRKOFFRAIMROEIGWFDVHVDGELNRLDQVSKINIGDKGMFFQSIATFTG 204
QY 187 FVVALASOKLALITMSVIPAEIVLTGICIAIDAAQEAERITRYSRAVLAEEVLSIRT 246
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 FIVGTPGKMLTVLAISPVLGSLAAIWAIIKLSFTDKELLAYAKAGAAEVLAAIRT 264
QY 247 VHAFYAGKKWVKYDVLQQAHOEKKKSPNNGVLPSTFECYVAAIALAFWKGFMYON 306
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 265 VIAFGGKKELEERYKNLEAKRIGIKKAITANISIGAAPLLIYASALAFWYGTSLVS 324
QY 307 GEVADYGVFTVTVLSTLAATSIHMLAPSGSVYQRRIFGSELSIDKPTQLDPLPSG 366
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 325 SEYT-IGQVLTFFSLGAFSGQSPSTEAFANARGAAYEIFKIIDKPSIDSYSKSG 383
QY 367 KQPEGCLGOIEONLAFYRPSRQAQVLRDNFTIPAGKTTALYGASGSGKSTVGLLER 426
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 HKPDNTKGNLEPKNFYHSRKEVKLKGLNKLQSGQTVALVNSGCGKSTTVQLMR 443
QY 427 WYLPSSGILLDLGELGYKVKLRIRLYVQOEPVLFRTGTFIQFIANGFEMDEQDLPRE 486
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 444 LYDPTDGMVCIQDQIRINVRHLREITGVVVSQEPVLFATTIENIRYV-----RE 494
QY 487 K-QMELVQKACKASNGDVFIMELNGYETEVGERAGALSGGQRIATARSISDPKILL 545
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 495 NVTMDIEKAVKEANAYDFIMKLPNKEDTLVGERGARLSGGQKQRIATARLVNPKILL 554
QY 546 LDEATSAIDPKAEKVVQALNRVSKDRTLVIAHKLATVKSAGNTAVISGKIVEGTHH 605
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 555 LDEATSAIDTESEAVVQVQALDKARKGRTIVIAHRLSTVRNADYTAGDDGVIVEKGNHD 614
QY 606 ELIEFCHYAALVRAODLGADEQOEHEKTLHEKAAAEAAERPALE-----RT 653
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 615 ELMKRKGITFKLVMTQTRGNEILEN-----ATGESKESDALEMSPKDSGSSLIKRR 667
```

Qy	917	STWFGRCG---PHLSQSMEFLATN---LGFCTADNDMTLOYSTTQOLANFISSSWAFCLPVQ	971
Db	936	SLRRHVGITFTSIQAMMYSTYAGCFRFGAIVLAGQIWEFQDVLVFS-----AYVEGAM	991
Qy	971	AAAQYLAYSTSTKARSANYILWLRLTKPTIRE-TEENKKKPGVGGCPVDLEDIEFPRYR	1021
Db	992	AVQGVSTFAPDYAKAKVSAAHINIEIKPLIDSYSTGLKPESTVEG-SVAFNDVFNYP	1051
Qy	1030	QRDSARVLGRGSMITPCQFVAVYVAGSGCGKSTLTALSFRFYDPTSGRISFAHENTAEIS	1081
Db	1051	TRPDVPLVRLGLSLVKKGQTLALVGGSGCGKSTVQVLLERFYDPLAGTVFDGSKVKOLN	1111
Qy	1090	PLRYRGHSLVQOEPITYLQGSVRENTYLALEAE-LSEELCGCLGPARPMLWILSLVOKA	1141
Db	1111	VQMLRAHMGVIOSEPLFDGCSIGENTAYGNSRVSSQSEIEHAAKEANIHSFIEMLDPKY	1171
Qy	1149	LKRLLAQRGMPGSGGORRTAIAARALIRNPKLLLLDEATSDALDTSERLVOAALDEASTS	1201
Db	1171	NTRV-GDKGTQLSGGQKORTAIAARALVROPHILLDEATSDALDTSERKVVQOALDKAREG	1221
Qy	1209	RTTIAVAHRLIRNVDFVIFANGRIAECTGTHAEIQRILGRYYEWMCIAQS	1259
Db	1230	RTCVIAHRLSTIQNALDVVFQNGRIKEGHTQQLLQAKGIYFTMVSQVA	1280
RESULT 12			
Qy	9JK64	PRELIMINARY; PRT; 1272 AA.	
ID	Q9JK64		
AC	Q9JK64		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DE	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN 1A.		
GN	PGY1		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
NCBI_Taxid	10116;		
[1]	SEQUENCE FROM N.A.		
RC	STRAIN=WISTAR; TISSUE=LIVER;		
RC	HOOVELD G.J.E.J.; WILMS J.W.J.; HAGENBUCH B.; JANSEN P.L.M.;		
RA	MEIJER D.K.F.; MULLER M.;		
RA	"Cloning and functional characterization of the rat multidrug		
RT	resistance protein Mdr1a";		
RL	Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL; AF257746; AAF69007.1; -.		
DR	InterPro; IPR003593; AAA.		
DR	InterPro; IPR001140; ABC_transporter_tmern.		
DR	InterPro; IPR003439; ABC_transporter.		
DR	InterPro; IPR001687; ATP_GTP_A.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
KW	ATP-binding; Transport.		
SQ	SEQUENCE 1272 AA; 140326 MW; BSCIDF89DFC986C5 CRC64;		
Query Match 25.4%; Score 1637.5; DB 11; Length 1272;			
Best Local similarity 32.3%; Pred. No. 2.2e-102;			
Matches 424; Conservative 254; Mismatches 532; Indels 101; Gaps			
Qy	3	DESEKPRPNODGSSSSHPPEKTEGSGISDYLIRIFRYADKYDWTNLNIALICAIIGSGAS	62
Db	4	EEDLNGRADKNFSKMGKSKKEKKPAVSVLTMERYAGWLDRFVMLLGLTALAIIHIGIA	63
Qy	63	LPLMSIIFSFTNKNFN-----YNSGDGSPFAKADVDHFLVFWFLYFGKFLVLY	113
Db	64	LPLMLVFGDMTDSFANVGNRRSMFSFYATDYAK-LEDEMTYATYITGAGVLLIVAY	122
Qy	114	VSTAATISAIARTTTTLRVLFRLECTLRQYVHFDKQSGAIAATXVTTNGNRIGTGAPEKL	173



QY 593 ISOGKIVEGTHETLIEGCHYAAVLRAQDLGADEQOEHE---KTLHEKAREANGERP 648  
 DB 593 FOGVIXEONGHDELREKGIYFLVMAQTAGNIEIENACEKSGDINDVMSKDSGS 652  
 QY 649 ALERTHTATIS-----QAGDLEKK-----VPVGLTGLSLKCLIMFYEOKNLYW-CF 696  
 DB 653 SLIRSTRKSRGPHDQGLSTKEALDDVPPASP-WRLKL-----NSTEWYF 703  
 QY 697 LLSTITVLCAATFFQOALLSRLTVPFELSG---HAAQERADFYILMFVWALGNLYG 754  
 DB 704 VVGVECAIINGGLQAFSIESKVVGVFTKNDTPETIQRONSNFSLFLILGILSIFTF 763  
 QY 755 TIGWTCNVISOVVTHYUQAAMFORVLDQDIELDIPQISGALTOLS-----802  
 DB 764 LQFTFGKAGEILTKRLMYPKSMRLQDISWDDPKNTGTALTFLANDAAQVKGAGS 823  
 QY 803 --ALPTQLOELISANFLIYVQHRLEQCSSTSLMETGPGCGVWCTSTPAFGWLPONS 860  
 DB 824 RLAVITONTANLGTGIIISLIY-----WQLTLLLLAIYV---ITAIAGVEMKMLSGQA 875  
 QY 861 SRDEARSLKGLCKRCWKACKSSYRDPDLIFDSRRPCSPVILGHVE---OGLA---- 912  
 DB 876 LKDKKELEGSGKIATEA-----IENFRTVVSILTREQKFETMAQGLQIPYR 921  
 QY 913 -KTIQSFWRGRCFGHLSQMEFLAIALGFCIAVDNWLQVSTTQLNFISSSWAFCLPVOA 971  
 DB 922 NALKKAHVGLTFSE--TQAMVFSYACFPFGAYLVARELMTFENVLLVFSIAIVFGAMA 979  
 QY 972 AAQYLAYSFTFKARSANYILWLTLPKPIRE--TEENKKGPVGGCPVDLEDIEFFRYQ 1030  
 DB 980 VGVSSFPADYAKAVYASHIRIEKPEIDSYSTGLKPNMCEG--WYKFNWGMFNTPT 1038  
 QY 1031 ROSARVLGVSMTEPIEGOFVAYVAGSGCGKSTLIALSERPYDTPSGRISFAHENIAENSP 1090  
 DB 1039 RPAPVPLQGLSLSEVKGOTLALVGSOGCGKSTVQVLLERFYDPMAGTVFLDGKEIKQJNV 1098  
 QY 1091 RLVRGMSLVQOEPTLYGSRVENVTLALEAEL--SEELCOGRLPARPMLTSLSLYOKA 1148  
 DB 1099 QWLRALHGLVSOEPLDFCSTAEINAYGDNSRVVSHIEIVKAKEAN--THQFIDSLPEKY 1157  
 QY 1149 LKRLAORGMOGSGGRORIARALRNPKLLLDLDEATSDLTQSERVQAALDEASTS 1208  
 DB 1158 NTRV-RDKGTQLSGGQKQRIARALVQPHILLDEATSDLTSEKVVQZALDKAREG 1216  
 QY 1209 RTTIANVHLRSLIRNDVIFVFAKRIETGTHAEQLORLGRYEMCLAQ 1259  
 DB 1217 RTCIVIAHLRSTIQNALDLIVIQNGQVKEGHTQPLAOKGIYFMSVQOA 1267  
 RESULT 14  
 ID 045721 PRELIMINARY; PRT; 1294 AA.  
 AC 045721; 062101;  
 DT 01-JUN-1998 (TREMELrel. 06. Created)  
 DT 01-NOV-1998 (TREMELrel. 08. Last sequence update)  
 DT 01-DEC-2001 (TREMELrel. 19. Last annotation update)  
 DE C47A10.1 PROTEIN.  
 GN C47A10.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Feloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN {1}  
 RP SEQUENCE FROM N.A.  
 RA Basmah V.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN {2}  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R.; Ainscough R.; Anderson K.; Baynes C.; Berks M.;  
 RA Bonfield J.; Burton J.; Connell M.; Copey T.; Cooper J.; Coulson A.;  
 RA Craxton M.; Dear S.; Du Z.; Durbin R.; Favell A.; Fulton L.

RA Gardner A.; Green P.; Hawkins T.; Hillier L.; Jier M.; Johnston L.;  
 RA Jones M.; Kershaw J.; Kirsten J.; Laister N.; Latreille P.;  
 RA Lightning J.; Lloyd C.; Murray A.; Mortimore B.; O'Callaghan M.;  
 RA Parsons J.; Percy C.; Rifkin L.; Roopra A.; Saunders B.; Shownkeen R.;  
 RA Smaildon N.; Smith A.; Sonhammer E.; Staden K.; Sulston J.;  
 RA Thierry-Mieg J.; Thomas K.; Vaubin M.; Vaughan K.; Waterston R.;  
 RA Watson A.; Weinstock L.; Wilkinson-Sproat J.; Wohlman P.;  
 RT "2.2 Mb-of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 RN {3}  
 RP SEQUENCE FROM N.A.  
 RA Basmah V.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -J- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL; Z93782; CAB07855.1; -;  
 DR EMBL; Z81484; CAB07855.1; JOINED.  
 DR EMBL; Z81484; CAB03973.1; -;  
 DR EMBL; Z93782; CAB03973.1; JOINED.  
 DR InterPro; IPR003593; AAA.  
 DR InterPro; IPR001140; ABC\_transporter\_tmem.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR001687; ATP\_GTP\_A.  
 DR InterPro; IPR000130; Zn\_Mtpeptidse.  
 DR Pfam; PF000664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 DR ATP-binding; Transport.  
 SQ SEQUENCE 1294 AA; 142490 MW; D8ED2602FA484DF CRC64;

Query Match 24.8%; Score 1601.5; DB 5; Length 1294;  
 Best Local Similarity 32.1%; Pred. No. 6.3e-100;  
 Matches 433; Conservative 234; Mismatches 509; Indels 175; Gaps 33;  
 QY 7 KPRPNQDGSSESSSH-----PPPEKETEGSIDYLRIFRYADKYDWTNLNIALICAGSGA 61  
 DB 5 KKKKDDSSSGSEKKEEAPPPPK-----ISIF-QLYRTSTVDRLMLAVGIIVSCATGV 58  
 QY 62 SLPLMSTIFGSGFTNKFNYY-----NSGDGSPFAFADVDHVF---LWFVYLFIGKPV 110  
 DB 59 GLPLMSILMGVNSQNPVTGLTFIDLPNSTASEKAAAEFSHEVQICNLYVVLGGIFA 118  
 QY 111 LTVYSTAAITISAIRTRRLRRVLECLTROYVHFKQSNKGAIAATVTTNGNRIOGTGIA 170  
 DB 119 AGFLQASCFMVICELNSNFRFQFFHVMRQEIAMDKNTSGTSLNKLFDNLNLRVBEGTG 178  
 QY 171 EKLIVFTVQALSMFSAFVVALASOMKALITMSVIPAFIVLTGICIA-IDAAOEARITRI 229  
 DB 179 DKVGLAFOMMAQFIGCFVAVFTYDMLLTLMNLSLP-FMMICGLFLAKLIATAATKEARQ 237  
 QY 230 YSRAAVLAEEVLSSIRTVHAFYAKKWKVKYDFVLOQAHOEGKKKSPNNGVLFSTFETFCI 289  
 DB 238 YAVAGGIAEEVLISIRTVAFNGQYECKREDALEHGKKTGKKSFLICAGLASFFVII 297  
 QY 290 YAAIALAFKGFNRYONGEADVGVKFTVVLSTLATSTISMLAPSGSVVYORRIFGSEL 349  
 DB 298 YASVCLAFWGTNFTVYSGRL-ESGTVLTFTVFMGSMGALGQAGQQQFATIGTALGAASL 356  
 QY 350 FSIIIDKPTQLDPLDPSGKQFEGCLQIEIQIALAFYPSRPSAOLVDFNLITIPAGKTAL 409  
 DB 357 YEVIDRIPEIDAYSTEGQTPSKISGRISVKNVKEEYTPTRADVKILKGVSLDAQPGQTV 416  
 QY 410 VGASGSGKSTMVGLLERWVLPSSGRILLDLGELQYVNVKRLRIRLVQGLVPLRGITF 469  
 DB 417 VGSSECKSTIIQLQRFYNPDAGQILDDIPEDFNKLYRLQVLGVVSGEPLNFTSIE 476  
 QY 470 QNTANGFMDQRDLPREKOMELVQKACKASNGDFINELPNGYETEVGEBACALSGQRQ 529  
 DB 477 QNIRYGRSDVSD-----EDIARALKEANAADFIKTFPEGLNTLVGORGVMGGQKQ 528



Db 646 QTESTDQKLSME-----SMKRSLSLRKSSLSRSLSKRSSSPSMFGFPAGIDTNNPAIP 699  
Qy 677 -----SLLKCILIMYE-----OKMLYWCFLLSITVILICAATFPQCALFLSRLLTVE 725  
Db 700 EKDIKVSTPIREKKVSFRVAALNKPEIPMLILGSIATAVLNGVILPIFGILISSVIAFF 759  
Qy 726 LSGHAAQERADYIILFFVVALGNLV-----YFTIGTCNVISQVVTTHRYOAAQFORV 779  
Db 760 KPPEQLKSDTRFWAIIIPMLGLVASVMVFFPAQTIFESIA-GCKLV-----QRIRSMCFEKV 813  
Qy 780 LDQDIELDIPQISGALTOLS-----ALPTQLOELISANF-LIYIVVGOHR 826  
Db 814 VRMEVGWFDETENSSGATGARLSADAATVRGLVGDAQAQTVQNLASVTAGLVIAFVASWQ 873  
Qy 827 LEQCSSTSLLWMTGPGCVCTSTPAFGMLPONSSR---DEARS-RKLGLKLCRCWACKR 882  
Db 874 LAFIVLAMLPL-IGLVYIYKFMVGFSADEASQVANDAVGSIKTVASFCAE----- 926  
Qy 883 SSYRDPDLIFDSRPGCPTVLGHVBOGLAKIIQSFWFCRCFGPHLSQSMEEUAIALGFC 942  
Db 927 -----ERVMKMYKKCEGPMRTGIRQIVSGI-----GFGVSEFFVLFSSYAASFY 971  
Qy 943 IA---VDNMLQVSTTQLNFISSSWAFCLPVQAAQVLAQYLAYSTFTKARSAANYILWLTLK 999  
Db 972 AGARLVDDG---KTFDSVFVFPALTMAVAISQSSLSUSDSSKASNAASIFAVIDRE 1028  
Qy 1000 PTIRETEENK-----KGPVGGCPVLEDEIEPRYQRDSARVLRGVSMTIEPGQFVAYVG 1054  
Db 1029 SKIDPSDESGRVLNVKG-----DIELRHISPKYPSRPDVQIFODLCLSTRAGKTIALVG 1083  
Qy 1055 ASGCGKSTLIALSERFYDPTSGRISFAHENIAEMSPRLYRGHMSLVQOEPTLYOGSVREN 1114  
Db 1084 ESGSGKSTVIALQRFYDPSGQITLDGVEIKTLQLKWLROQTGLVSOEPVLFNETIRAN 1143  
Qy 1115 VTL-----ALEAEL--SEELCOGRUPARPMWLTLSSLYQALKRLAAQRMQFSGGQOR 1167  
Db 1144 IAYKGGGDATETEIVSAAELSNH-----GFISGL-QQGYDTMVGERGVOLSGGQOR 1195  
Qy 1168 IATARALIRNPKLLLDLDEATSDLTOSERLVQAALDEASTRTTIAVAHRLSTIRNVDVI 1227  
Db 1196 VATAIRVQKPKVLLLDLDEATSDLAESERVVQDALDRVMVNETTVVVAHRLSTIKNADVI 1255  
Qy 1228 FVFANGRIATGTHAELQRLRGYY 1252  
Db 1256 AVYKNGVIVEKGHETLINIKDGYI 1280





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 07:59:41 ; Search time 35.22 Seconds

(without alignments)  
1388.495 Million cell updates/sec

Title: US-09-882-694A-11

Perfect score: 6455

Sequence: 1 MADESEKPRNPQDSSESSH.....LORLRGRYVMCLAQSLDOA 1263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1970.5	30.5	1362	1 PMD1_SCHPO	P36619 schizosacch
2	1649.5	25.6	1280	1 MDR1_HUMAN	P08183 homo sapien
3	1640.5	25.4	1279	1 MDR3_HUMAN	P21439 homo sapien
4	1630.5	25.3	1276	1 MDR1_MOUSE	P06795 mus musculus
5	1627.5	25.2	1276	1 MDR3_MOUSE	P21447 mus musculus
6	1619.5	25.1	1276	1 MDR2_CRIGR	P21449 cricetus
7	1612	25.0	1276	1 MDR1_CRIGR	P21448 cricetus
8	1601.5	24.8	1276	1 MDR2_MOUSE	P21440 mus musculus
9	1592.5	24.7	1278	1 MDR2_RAT	P08201 rattus norv
10	1590.5	24.6	1281	1 MDR3_CRIGR	P23174 cricetus
11	1590	24.6	1277	1 MDR1_RAT	P43245 rattus norv
12	1551	24.0	1321	1 MDR1_CAEEL	P34712 caenorhabdi
13	1507.5	23.4	1302	1 MDR4_DROME	Q00449 drosophila
14	1472.5	22.8	1321	1 AB11_RAT	O70127 rattus norv
15	1468.5	22.7	1321	1 AB11_MOUSE	O70127 mus musculus
16	1463	22.7	1321	1 AB11_HUMAN	O95342 homo sapien
17	1453	22.5	1254	1 MDR3_CAEEL	P34713 caenorhabdi
18	1406.5	21.8	1321	1 AB11_RABIT	Q9u0v3 oryctolagus
19	1403.5	21.7	1302	1 MDR5_DROME	Q00748 drosophila
20	1286	19.9	1280	1 MDR1_LEIEN	Q06034 leishmania
21	1011.5	15.7	1419	1 MDR_PLAFA	P13568 plasmodium
22	832.5	12.9	1290	1 STE6_YEAST	P12866 saccharomyc
23	814	12.6	1336	1 MAM1_SCHPO	P78966 schizosacch
24	713.5	11.1	735	1 ABC8_HUMAN	O9nu22 homo sapien
25	690	10.7	582	1 MSBA_ECOLI	P27259 escherichia
26	641	9.9	1325	1 MDR4_HUMAN	O15439 homo sapien
27	640	9.9	726	1 YFX9_SCHPO	Q9y7m7 schizosacch
28	640	9.9	820	1 MDL2_YEAST	P33311 saccharomyc
29	631.5	9.8	1436	1 MRP5_RAT	O9qym0 rattus norv
30	624.5	9.7	1502	1 MRP6_RAT	O88269 rattus norv
31	623.5	9.7	1503	1 MRP6_HUMAN	O95255 homo sapien
32	611.5	9.5	725	1 TAP1_RAT	P36370 rattus norv
33	610.5	9.5	1436	1 MRP5_MOUSE	Q9rlx5 mus musculus

34	610	9.5	724	1 TAP1_MOUSE	P21958 mus musculus
35	607.5	9.4	703	1 TAP2_RAT	P36372 rattus norv
36	605	9.4	1437	1 MRP5_HUMAN	O15440 homo sapien
37	603.5	9.3	583	1 EXP8_STRPN	P35598 streptococc
38	601.5	9.3	1581	1 ACC8_RAT	O09429 rattus norv
39	597.5	9.3	1323	1 HS76_CANAL	P53706 candida alb
40	595.5	9.2	587	1 MSBA_HAEIN	P44407 haemophilus
41	595	9.2	1548	1 MDR_LEITA	P21441 leishmania
42	594.5	9.2	1531	1 MRP1_HUMAN	P36327 homo sapien
43	591.5	9.2	702	1 TAP2_MOUSE	P09427 mus musculus
44	591.5	9.2	1581	1 ACC8_CRIGR	O09427 cricetus cr
45	589	9.1	631	1 YC72_MYCTU	Q11047 mycobacteri

## ALIGNMENTS

RESULT 1					
PMD1_SCHPO					
ID	PMD1_SCHPO	STANDARD;	PRT;	1362 AA.	
AC	P36619; 074513;				
DT	01-JUN-1994 (Rel. 29, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Leptomycin B resistance protein pm1.				
GN	PMD1 OR SPCC663.03.				
OS	Schizosaccharomyces pombe (fission yeast).				
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;				
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;				
OC	Schizosaccharomycetes.				
OX	NCBI_TaxID=4896;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92244046; PubMed=1349418;				
RA	Nishi K., Yoshida M., Nishimura M., Nishikawa M., Nishiyama M.,				
RA	Horinouchi S., Beppu T.;				
RT	"A leptomycin B resistance gene of Schizosaccharomyces pombe encodes				
RT	a protein similar to the mammalian p-glycoproteins.";				
RL	Mol. Microbiol. 6:761-769(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=972;				
RA	Seeger K., Harris D., Wood V., Rajandream M.A., Barrell B.G.;				
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: MAY BE A TRANSMEMBRANE TRANSPORTER OF THE MATING				
CC	FACTOR, NAMELY P-FACTOR OR M-FACTOR. CONFERS RESISTANCE TO				
CC	LEPTOMYCIN B AND TO SEVERAL OTHER ANTIFUNGAL DRUGS.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).				
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; D10695; BAA01537.1; ..				
DR	EMBL; AL031307; CAA20363.1; ..				
DR	PIR; S20548; S20548.				
DR	InterPro; IPR003593; AAA.				
DR	InterPro; IPR001140; ABC transporter_tmem.				
DR	InterPro; IPR003439; ABC transportr.				
DR	InterPro; IPR001687; ATP_GTP_A.				
DR	Pfam; PF00664; ABC_membrane; 2.				
DR	Pfam; PF00005; ABC_tran; 2.				
DR	SMART; SM00382; AAA; 2.				
DR	PROSITE; PS0211; ABC_TRANSPORTER; 2.				
DR	Transmembrane; ATP-binding; Glycoprotein; Antibiotic resistance.				
KW	DOMAIN 91				
FT	TRANSMEM 122				
FT	POTENTIAL.				
FT	TRANSMEM 91				
FT	POTENTIAL.				
FT	TRANSMEM 138				
FT	POTENTIAL.				

FT TRANSNM 220 237 POTENTIAL.  
 FT TRANSNM 244 264 POTENTIAL.  
 FT TRANSNM 320 346 POTENTIAL.  
 FT TRANSNM 354 374 POTENTIAL.  
 FT DOMAIN 375 788 CYTOPLASMIC (POTENTIAL).  
 FT TRANSNM 789 809 POTENTIAL.  
 FT TRANSNM 835 859 POTENTIAL.  
 FT TRANSNM 916 935 POTENTIAL.  
 FT TRANSNM 940 957 POTENTIAL.  
 FT TRANSNM 1022 1040 POTENTIAL.  
 FT TRANSNM 1054 1072 POTENTIAL.  
 FT DOMAIN 1073 1362 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 455 462 ATP (POTENTIAL).  
 FT NP\_BIND 1134 1161 ATP (POTENTIAL).  
 FT CONFLICT 1334 1335 IA -> TC (IN REF. 1).  
 SO SEQUENCE 1362 AA: 149631 MW: 2ADF8AEB7B3ACE8 CRC64;

Query Match 30.5%; Score 1970.5; DB 1; Length 1362;  
 Best Local Similarity 35.5%; Pred. No. 7.4e-120;  
 Matches 484; Conservative 225; Mismatches 508; Indels 147; Gaps 24;

QY 9 RPNDSGES-----SSHPPEKE-----TEGSISDYLRIFRYADKYDNTLVY 50  
 DB KSENEASESHVVVDKDFEQYTPPEQILYKQINDPAKLSGYPRILLSTADKAWDMLQL 96  
 QY 51 IALICAGSGLPLMSIFGFTNNKPNYNSGDSPKADVDHFLVFLYFTGKRV 110  
 DB 97 AGTITGIGAGLPLMSVLSQLOAQAFDLASGRGA--SSFOHTVDHECLFYIYIAIGVFG 155  
 QY 111 LTVYSTAAITISAITRTLRVLECYLRQEVNHFDPKQNGALATATVTNGNRIGTGA 170  
 DB 156 CSYIYVTELIAGERIARRIQDYHLAILSONIGYFDRLAGAGEITRTITDTNFIDQLG 215  
 QY 171 EKLVTQVQALSFFAFVVALASOKKALITMSVIPALFVLTGICINADAAQARITRY 230  
 DB 216 EKVLVFAIATVSGVFIATIRHKKFTLILSSMFPALCGGILGVFPFTKNTKGQIAVV 275  
 QY 231 SRAVLAEVLSISRTVIAFYAQKWKVEKYDVFLOAHQSGKKSPNNGVLESTVEFY 290  
 DB 276 AESSTVEEVFNIRNAFAFGTDILAKLYNKYLITAQREFGINKAKAMGLMVGWMEFVAY 335  
 QY 291 AATLAFKFRMYONGEVADYGVVTVVLSVTLAATISISMLAPSGSVVYVORRIFGSELF 350  
 DB 336 GYVGLAFWEGGRLLIAGDL-DVSKLIGCFFAVLIATSYLANISPKMQSFVSCASAKKIF 394  
 QY 351 SIIDKPTOLDPLDPSGKPEGLQIEIONLAFAYPSRPSAQVLRDENTLIPAGKTTALV 410  
 DB 395 DTIDRVSPINAPPTGDDVVKIKGEIELKNIRFVYPTPEVLVDNDFSLVCPGSKITLV 454  
 QY 411 GASGSKSTWGLLRLWLPSSGRILLDGLGQYNNVKNLRSIRLVQOQSPVLPFGTIFQ 470  
 DB 455 GASGSKSTIIGLVERFYDPTGGQVFLDGLDKDLRTLVASLRNOISLVQOQEPVLPFATVPE 514  
 QY 471 NTANGPMDF-ORDLPRKQMLVOKACKASNGDVFINELPNGYETEYGERAGALSGGQ 529  
 DB 515 NITYGLPDTIKGTLSEELRRVYDAKLANAYDFIMTLPEQFSTNVGQGLMSGGQK 574  
 QY 530 RTAARSISDPKILLDEATSDAPKAEKVQVQALNRVSKDRTTLVIAHKLATVKSAGN 589  
 DB 575 RTAARAVISDPKILLDEATSDALDSKSEVLVQKALDNASRSTTVIAHRLSTIRNADN 634  
 QY 590 IAVISGKIVGQTHHEILIEFGCHYALVRAQDL-GADEQOQHEKTLHKAARE----- 642  
 DB 635 IVVYNAKIVGQSNELLIDNGAYALVEAQKLSGGEKDOEMVEELEDAPREIPITSF 694  
 QY 643 -----AAGERPALERTHTTATSQAGLEKRR-----VP----- 670  
 DB 695 GODDEDNDMASLEAPM--SHNTDITDNLNKLNEKNVVFEDKTLQHVASEIVNLPAPD 752  
 QY 671 VGTGLG-----YSLKLCILIMEFYQKKNLYWCFLISTVYLI 705  
 DB 753 VQELNEEPKSKSKKNHNEINSLTALWFIHSFVTHIEII-----CLLIGILASHI 804

QY 706 CAATFEGQALLFRLLTVF-ELSGHAAOERADFYILMFVVALGNLGVYFTIGWTCNVIS 764  
 DB 805 CGAAYPQAAVAFARFLNIETDLSSDFLHKVNVFAYVYLLIATVQFFAYATSNFAMTYAM 864  
 QY 765 QVYTHRYQAAAFORVLDQDIELDIPQEQISSALTSQALPQLOEL-----ISANFL 817  
 DB 865 EAVLQRIYHLFRTLRLQDVEFFDRSENVTGAITTSLS---TKIQLRGLSGPTLGTFFQ 921  
 QY 818 IYIVVGOHRLQEQSTSLMNETGPGCVWCTSTPAFGWLPQNSSRDEARSRLKLCRK - 876  
 DB 922 ILTNI-----ISVTILSLAYG-----MKLGLVLTSTSPVITAGYVYRVALDQVQEK 969  
 QY 977 -----CWAKRRSSYRDPDLIFDSRRP-----CSPTVLGHVGEGLAKIITQSFWFG 921  
 DB 970 SAAYKESAAFAFACES---AIRTVASLNREENVFAEYCDLSIKPGRESAIALSKSGLFFS 1026  
 QY 922 RCEGFHLSQSMFLAIALGFCIAVDNMLQVSTQTNFISSSWAFCLPVQAAAAYLAYSTS 981  
 DB 1027 -----AAQGVTELINALTFTWYSTLMRKGEYNIQVFTCFIAIVFGIQQAQGFYSAD 1080  
 QY 982 FTKARSANVILRLTKPTIRE-TEENKKGKPGVGCPCVDELEDFRYRORDSARLVGV 1040  
 DB 1081 VTKARAAAGKYLSESKPIDTSTWTEGKKVLSQSAIEFROVEFYPTRRHKKVLRLG 1140  
 QY 1041 SMTIEPQGVVAVGAGCGCKSTLIALSERFYDPTSGRISFAHENAEMSPRLYRGHMSLV 1100  
 DB 1141 NLTVKPGQFVAFVYSSGCGKSTTIGLIERFYDCDNGAVLVDGVNVYRDIYNDYRKQIALV 1200  
 QY 1101 QOEPYLYQGSVRENVTLALEAELS-EELCQGRLPARPMWLILSSLYQKALKRLAAQRMQ 1159  
 DB 1201 SQEPYLYQGVRENVTLGASKDVSEEMIEACKKANIHFILG--LPNGYNTLCGQSS 1258  
 QY 1160 FSGQORTAIPALRNPKLLLDDEATSDALDQSERLVOALDEASTSTRTTIAVAHRLS 1219  
 DB 1259 LSGQORTAIPALRNPKLLLDDEATSDALDQSERLVOALDEASTSTRTTIAVAHRLS 1318  
 QY 1220 TIRNVDFVIFANGRIATGTHAELQRLRGYRYEMCLAOGLDQA 1263  
 DB 1319 SIQDAPCFYFDGVIAGAEAGTHAELVQGRYVELVVEQGLNKA 1362

RESULT 2  
 ID MDRL\_HUMAN STANDARD; PRT: 1280 AA.  
 AC P08183; Q12755; Q14812;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Multidrug resistance protein 1 (p-glycoprotein 1).  
 GN ABCB1 OR PGP1 OR MDRL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87028230; PubMed=2876781;  
 RA Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,  
 RA Roninson I.B.;  
 RT "Internal duplication and homology with bacterial transport proteins  
 RT in the mdrl (p-glycoprotein) gene from multidrug-resistant human  
 RT cells.";  
 RL Cell 47:381-389(1986).  
 RN 2;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90094448; PubMed=1967175;  
 RA Chen C.-J., Clark D.P., Ueda K., Pastan I., Gottesman M.M.,  
 RA Roninson I.B.;  
 RT "Genomic organization of the human multidrug resistance (MDR1) gene  
 RT and origin of p-glycoproteins";  
 RL J Biol. Chem. 265:506-514(1990).  
 RN 3;



Db 445 DPTEGMVSDGDDIRTINVRFLREILIGVVSQEPVLPATTIAENIRYG-----RENV 495  
QY 488 OMELVOKACKASGVDFINELPNCYETEVCEAGALSGGQORORATARSIIISDKPILLD 547  
Db 496 TWDEIEKAVKANAYDFIMKLPKFDVLVGERGAQLSGGQORAIARALVNRPKILLD 555  
QY 548 EATSALDPKAKVVOEALNRVSDRTLVIAHKLATVKSAGNIAVISGKIVGQTHHEL 607  
Db 556 EATSALDESAVQVQALDKARGRRTVIAHRLSTVRNADVIAGFDGVDGIVVERGNHDEL 615  
QY 608 IEFCHYALVRAODLGADEQREKTLHEKAAREAAERPALERT----- 653  
Db 616 MREKGIYFKLVMTAGNEVELEN-----AADESKSEIDALESSNDSSRLIRKRST 668  
QY 654 -HTTATSQAGD-----LEKRVKVPVGTGLYSLLKILIMPEQKNLYW-CFLLSTIV 703  
Db 669 RRSVGSQAQDKLSTREALDESIPVPSF--NRIMKLN-----TEMFYFVGVFECA 718  
QY 704 LICAATFPQALLESRLITVPEL--SCHAAQERADPFYILMPFVALGNLVGYFTIGTCN 761  
Db 719 IINGLOPAPALIESKIIIGVTRIDDPETKRQNSNLSLFLALGIIIFITFFLOGFTG 778  
QY 762 VTSQVVTTRYQAAMFORVLDIELDIPETISGALTSOLSALPTQLQELISANFLIYV 821  
Db 779 KAGEILTTRKRYMYFRSMLKQDSWFDKNTGTGALTTRIANDAQAQVKAIGSLAV--- 835  
QY 822 VQHRLEQCTTSLMMETPGGCVNCTSTPAFGW-----LQ 858  
Db 836 -----ITONIANLGTGII-----SFYGMQLTLLLAIVPIIAIAGVEMKMLSG 881  
QY 859 NSSRDEARSRLKGLCKRCKWACKRSYRDPDRLIPDSRRPCSTVLGHVGEQGLAKIIO-- 916  
Db 882 QALDKKKEGAGKATEA-----IENFTVYSLTQEQKFEHMYAQSLQVP 927  
QY 917 -----SFWRGCFGPHLSQMEFLAIA-----LGFCIAYDNWLVQVSTQLNFSSWAF 965  
Db 928 YNSLRKAHIGITFSF--TOAMHFSYACFRFGALYVAHKLMSEFVLLVS-----AV 981  
QY 966 CLPQQAAYLAKYSTTKARSANILMLRLKPTI--REFEENKKGPVGGCPVLE- 1022  
Db 982 VEGAMAVGVSSAPDPAKAKISAHIIMIETPLIDISYFE-----GLMPNTLEG 1033  
QY 1023 -----DIFRYRDSARVLGRVSMITPFGQFVAVVGASGCKSTLIALSERFYDPTSGR 1077  
Db 1034 NTFGEVVFNTPTPDIPLVLOGLSLEVKKGOTLALVSGGCKSTVQLLRFYDPLAG 1093  
QY 1078 ISFAHENIAEMSPRLYGHMSLVQEPITYOGSVRENVTLEAEL--SEELCOGRLPAR 1135  
Db 1094 VLLDGKEIKRLNVQWLAHLGIVSQEPILFDCSIAENIAYGDSRVVSQEEIVRAAKEAN 1153  
QY 1136 PMLWTLSSLYOKALKRLAAGRMQFSGGQORATARALIRPKILLDEATSLDTSO 1195  
Db 1154 IHAFIESLPNKYSTK--VGDKGTLGSGGQORATARALVQPHILLDEATSLDTSO 1211  
QY 1196 RLVQALDEASTRTTAVAHRLSTTRNDVDFVFNANGRIAPTGHAEQLRLGRYEMC 1255  
Db 1212 KYVQALDKAREGRTCVIAHRLSTIONADLIYVFGVRKHGPHQOOLLAQKGIYFSMV 1271  
QY 1256 LAQS 1259  
Db 1272 SVQA 1275

RESULT 3  
MDR3 HUMAN  
ID MDR3 HUMAN STANDARD; PRT: 1279 AA.  
AC P21439;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Multidrug resistance protein 3 (P-glycoprotein 3).  
GN ABCB4 OR PGY3 OR MDR3.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN NCB TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=89138016; PubMed=2906314;  
RX van der Bliek A.M., Koolman P.M., Schneider C., Borst P.;  
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";  
RL Gene 71:401-411(1988).  
[2]  
RN SEQUENCE OF 856-1279 FROM N.A.  
RX MEDLINE=8811519; PubMed=2892668;  
RA van der Bliek A.M., Baas F., ten Houte de Lange T., Koolman P.M.,  
RT van der Velde-Koerts T., Borst P.;  
RL "The human mdr3 gene encodes a novel P-glycoprotein homologue and  
gives rise to alternatively spliced mRNAs in liver.";  
RL EMB J. 6:3325-3331(1987).  
[3]  
RN GENE STRUCTURE  
RX MEDLINE=91161629; PubMed=2002063;  
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;  
RT "Structure of the human MDR3 gene and physical mapping of the human  
MDR locus.";  
RL J. Biol. Chem. 266:5303-5310(1991).  
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT  
CC CAPABLE OF CONFERRING DRUG RESISTANCE. MEDIATES THE TRANSLOCATION  
CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE  
CC HEPATOCYTE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- DISEASE: DEFECTS IN ABCB4 ARE A CAUSE OF PROGRESSIVE FAMILIAL  
CC INTRAHEPATIC CHOLESTASIS TYPE III (PFIC), A FORM OF AUTOSOMAL  
CC RECESSIVE LIVER DISORDERS, CHARACTERIZED BY EARLY ONSET OF  
CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE  
CC ADULTHOOD.  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M23234; AAA36207.1; -;  
DR EMBL: X06181; CAA29547.1; -;  
DR PIR: JS0051; DVH03.  
DR PIR: A42213; A42213.  
DR HSP: P13569; INBD.  
DR MIM: 171060; -;  
DR MIM: 602347; -;  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_Gtp\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
KW Multigene family.  
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 58 78  
FT TRANSMEM 123 143 POTENTIAL.  
FT TRANSMEM 192 211 POTENTIAL.  
FT TRANSMEM 216 235 POTENTIAL.  
FT TRANSMEM 301 320 POTENTIAL.  
FT TRANSMEM 336 354 POTENTIAL.  
FT DOMAIN 355 711 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 712 732 POTENTIAL.  
FT TRANSMEM 756 776 POTENTIAL.  
FT TRANSMEM 832 851

FT TRANSMEM 854 873 POTENTIAL.  
FT TRANSMEM 937 956 POTENTIAL.  
FT TRANSMEM 976 993 POTENTIAL.  
FT DOMAIN 994 1279 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 429 436 ATP (BY SIMILARITY).  
FT NP\_BIND 1069 1076 ATP (BY SIMILARITY).  
FT REPEAT 641 640  
FT REPEAT 641 1279  
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 1093 1093 V -> VFVDFGFO (IN REF. 2).  
SQ SEQUENCE 1279 AA; 140682 MW; 3058C9B5C8D6087 CRC64;

Query Match 25.4%; Score 1640.5; DB 1; Length 1279;  
Best Local Similarity 33.3%; Pred. No. 1.8e-98;  
Matches 422; Conservative 230; Mismatches 532; Indels 89; Gaps 21;

QY 35 LRIFRYADKYDWTNLNLTALICAGSGASLPLMSIIFGSPNKNFNYSNDSGSPFAF---- 90  
DB 43 LTLFRYSQDKLWMSLGTIMAHGSLPLMIVFEGMTDKPVDTAGNFPVNFSL 102  
QY 91 -----KADVDHFVLMFVYLFIGKFLVTVVSTAAITISAIKTRTFLRRVFLLECTLRQEV 143  
DB 103 LNPGRILEEMTRAYIYSGLAGLVAAVIQVSWFTLAAGROIKRQKPPHAILRQEI 162  
QY 144 WHFDKOSGALATVVTNGNRIGTQIAEKLVFTVQALSMEFFSAFVVVALASOMKALITMS 203  
DB 163 GWFDINDITELNTRLTDISKISGIGDKVGMFFQAVATFFAGFVIGFIRGKKLTVLMA 222  
QY 204 VPAFLVGTGICIAIDAOEARTIRYSRAVLAEVLSIRTVHAFYAQKKWKEVDVF 263  
DB 223 ISPIILGSAAYWAKILSAFSDKELAAAYAKAGVAEALGAIRTVTAFGONKELERYOKH 282  
QY 264 LOOAHQEGKKSPNNGVLFSTEXFYCAIATALAFKGRMYQNGEVADYGVFTTVVLSVT 323  
DB 283 LENAKEIGIKKAISANISMGIAFLIYASVALAFYGTSLVISKEYT-IGNAMTVFESIL 341  
QY 324 LAATSISMLAPSGSVYORRIFGSELSIIDKPTOLDPLDPSGKQEGCLGOIETONLAF 383  
DB 342 IGAFSVGOAACPIDAFANARGAANVIFDIIDNPKIDSERGHKDSIKGNLEFNDFVF 401  
QY 384 ATPSPSAQVLBNFLIPAGKTALVAGSGCKSTWGLLERWLPSSGRILLDGLLEG 443  
DB 402 SYPSRANKILKLNKLVQSGQTVLVGSSGCKSTVQLIQRLYDDEGTINIDQDIR 461  
QY 444 QYNVWLSRLVQOQEPVLPFRGTIFQNIANG-----FMDEQDLPREKOMELVOKACKAS 499  
DB 462 NFNVYLREILGVVSQEPVLPFTTIAENICYGRGNVTMDE-----IKKAVKEA 509  
QY 500 NGDVFTELNPNGYETEGERAGALSGGORORIAARSIIIDPKILLDEATSAIDPKAEK 559  
DB 510 NAYEFIMKLPQKFDTLVGERAQLSGGCKORAIARALVYNPKILLDEATSAIDTSEFA 569  
QY 560 VVOEALNRVSKDRTLVIAHKLATVKSAGNIIVISOGKIVEQGTTHHELFEPCCHVALVR 619  
DB 570 EYQAALDKAREGRTIVIAHLSLTVRNADVIAGFEDGVIVEQSHSELKKEGVTFKLVN 629  
QY 620 AODIGADEQOEHKTLHEKAREAA--GERPALERTHTTATSOAGLEKRRKVPVGTG-- 675  
DB 630 MOTSGSQIOSEBEFELNDEKAATMAPNGKSLFRHSTQKNLNSOMCOKSLDVTGDLG 689  
QY 676 -----YSLKCTILIMFYEQKKNLYWCFLITITVILICATFPQOALLFLSLLTVFELS 730  
DB 690 ANVPVSEFLVKL-----NKTWEPYFVVGTVCAIANGGLQPAFVSIFSLIIAIFGDDA 745  
QY 731 A--OERADYILMFFVALGNLVGFTTGTWCNVISQVWHRVYQAAQFQVILQDLELDI 789  
DB 746 VQOQKCNIESLIFLGLIISFFTEFQGTFFGKAGEILTRRLRSMAFKAMLRQDSWDFD 805  
QY 790 PEQISQALTSQLSALPTQLO-----ELISAN-----FLIYIVVGQHRLEQCSSTLSL 835  
DB 806 HKNSTGALSTRLADAAQVOCATCTRLALIAQNIANLGTGIIISFIYG-----NQLTIL 859

QY 836 WMETGPGGCVMTCTSPAFGWLPONSRDSEARSKLGLKCRKCKWACKRSYRDPDRLIFDS 895  
DB 860 LLAVVP--IIAVGIVEMKLLAGNAKRDKELEAGKIAATEAENIRTVSLVTQERKFS 917  
QY 896 RRPCSPTVLGHVVEOGLAKIIOSFWFGRCFGFHLQSMEFLAIALGPCIAVDNWLQVSTTO 955  
DB 918 M-----YVEKLYGPIRNSVQKAHIYGITFSISQAFWYFVYACGFRGAYLVN-GH 967  
QY 956 LNFISSSNAFCLPVQAAQYAYSTSP-----TKARSANILMLTLTKPTIRE-TEENKK 1010  
DB 968 MRFVDVILVFSIAIVGAVA-LGHASSAPQYAKAKLSAHLFMLFERQPLIDSSEGLK 1026  
QY 1011 KGPVGGCPVDLEJIEFRYQRDSARVLGRVSMITTEPOQVAYVYAGSGCGKSTLIALSERF 1070  
DB 1027 PKFEG-NITFNEVVENPTRANVPVQLGSLVKGOTLALVSSGCGKSTVVQLLERF 1085  
QY 1071 YDPTSGRISFAHENAEMSPRLYRCHMSLVQOEPVLYOGSVRENVTLALAEEL--SEELC 1128  
DB 1086 YDLAGTVLLOQGEAKKLVQWLRQAQLGIVSQEPIFLDCSIAENIAYGDSRVVSQDEIV 1145  
QY 1129 QGRLEPA--RPMILWISSLYQKALKRLAAORQWFGSGGQRORIAARALINPKLLLDEA 1186  
DB 1146 SAKAANIHPFTIETLPHRYETRV---GDKGTOLSGGQKORIAARALIPQOILLDEA 1201  
QY 1187 TSALDTQSERLVOAALDEASTSRTTIAVAHRLSTIRVNDVIFVFANGRIAEETGTHAELQR 1246  
DB 1202 TSALDTSEKVVQOALDKAREGRTICIVIAHRLSTIQNALDILVVFQNGRYKHEGTHOQLLA 1261  
QY 1247 LAGRYEMCLAQS 1259  
DB 1262 QKGIFYSWMSVOA 1274

RESULT 4  
MDR1\_MOUSE  
ID MDR1\_MOUSE STANDARD; PRT; 1276 AA.  
AC P06795;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Multidrug resistance protein 1 (P-glycoprotein 1).  
GN ABC1 OR PG11 OR PG11-1 OR MDR1 OR MDR1B.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87028229; PubMed=3768958;  
RA Gros P., Croop J., Housman D.;  
RT "Mammalian multidrug resistance gene: complete cDNA sequence  
RT indicates strong homology to bacterial transport proteins.";  
RL Cell 47:371-380(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89367274; PubMed=2570420;  
RA Raymond M., Gros P.;  
RT "Mammalian multidrug-resistance gene: correlation of exon  
RT organization with structural domains and duplication of an ancestral  
RT gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).  
RN [3]  
RP SEQUENCE OF 1-21 FROM N.A.  
RX MEDLINE=91042535; PubMed=2248681;  
RA Raymond M., Gros P.;  
RT "Cell-specific activity of cis-acting regulatory elements in the  
RL promoter of the mouse multidrug resistance gene mdr1.";  
RL Mol. Cell. Biol. 10:6036-6040(1990).  
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE

CC CC LINKER DOMAIN.  
CC CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE  
CC CC RELATED BUT DISTINCT CELLULAR GENES.  
CC CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.  
CC CC -----  
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC CC the European Bioinformatics Institute. There are no restrictions on its  
CC CC use by non-profit institutions as long as its content is in no way  
CC CC modified and this statement is not removed. Usage by and for commercial  
CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC CC -----  
CC CC EMBL; M14757; AAA79005.1; -;  
CC CC PIR; A33719; DVMSL  
CC CC MGI; 97568; Abcd1.  
CC CC InterPro: IPR003593; AAA.  
CC CC InterPro: IPR001140; ABC\_transporter\_tmem.  
CC CC InterPro: IPR003439; ABC\_transporter.  
CC CC Pfam; PF00664; ABC\_membrane; 2.  
CC CC Pfam; PF00005; ABC\_tran; 2.  
CC CC SMART; SM00382; AAA; 2.  
CC CC PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
CC CC APP-binding; glycoprotein; Transmembrane; Transport; Repeat;  
CC CC Multigene family; Phosphorylation.  
CC CC DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).  
CC CC TRANSMEM 48 71 POTENTIAL.  
CC CC TRANSMEM 119 139 POTENTIAL.  
CC CC TRANSMEM 199 219 POTENTIAL.  
CC CC TRANSMEM 296 316 POTENTIAL.  
CC CC TRANSMEM 327 347 POTENTIAL.  
CC CC TRANSMEM 348 709 POTENTIAL.  
CC CC TRANSMEM 710 730 POTENTIAL.  
CC CC TRANSMEM 755 775 POTENTIAL.  
CC CC TRANSMEM 831 851 POTENTIAL.  
CC CC TRANSMEM 855 882 POTENTIAL.  
CC CC TRANSMEM 935 955 POTENTIAL.  
CC CC TRANSMEM 966 986 POTENTIAL.  
CC CC DOMAIN 987 1276 CYTOPLASMIC (POTENTIAL).  
CC CC NP\_BIND 426 433 ATP (BY SIMILARITY).  
CC CC NP\_BIND 1068 1075 ATP (BY SIMILARITY).  
CC CC REPEAT 1 635  
CC CC REPEAT 636 1276  
CC CC CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CC CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CC CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CC CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CC SEQUENCE 1276 AA; 140993 MW; 18040011B0FFAE CRC64;

Query Match 25.3%; Score 1630.5; DB 1; Length 1276;  
Best Local Similarity 31.6%; Pred. No. 7.8e-98;  
Matches 422; Conservative 250; Mismatches 517; Indels 145; Gaps 22;  
QY 3 DESEKPRNQDSSESSHPPEKETEGSISDYLRIFRYADKYDWTNLNIALICAIGSGAS 62  
DB 4 EENLGRADNFKSMGKSKREKKEKPAVGFGFRYADWLKMLGTLTAIRIHGTL 63  
QY 63 LPLMSIFGSFTNKN-----WYNSGDGSP-----EAFKADVDHFLVFLVFLFTG 107  
DB 64 LPLMLVFGNMTDSFTKAEASILPSITNQSGPNSTLISSSLEEEEMAIYAYTYTGIGAG 123  
QY 108 KFLVTYVSTAATISAIKTRTRLRVLECFTRQVWFHFKQSGATATVTTNGNRIO 167  
DB 124 VLIVAIQYSLWCLAAGRIHQKFFHATMNGEIGWFDVHDVDELNTRLTDVSKIND 183  
QY 168 GIAEKLVFTVQALSWFFSAFVALASOKKALITMSVIPALFTVGTICIAIDAQAARIT 227  
DB 184 GIGDKICMFQSIYTFIAGPIGFSIGWKLTVILVSPILGSSALWAKVLTSFTNKL 243  
QY 228 RIYSRAVLAELVLSIRTVHAFYAQKKWKEKYDVLQQAHEGCKKSPNNGVLFSTEYP 287

DB 244 QAYAKAGAAVEEVLAAIRTVIAFGGQOKELEKRYNKNLEAKNVGKIKKAITASIGAYL 303  
QY 288 CIYAAIALAFWKGFMYONCEVADYKVFYVTVLSVTLAATSISMLAPSGSVVYORRIFGS 347  
DB 304 LVYAVLAFWYGTSLVLSNEYS-IGEVLTFFSILGTFSGHILAPNIEAFANARAAAF 362  
QY 348 ELESIDKPTQDLPDPSKOPGCGLGOIEIQNLAFAYPSRPSAQVLRDFNLTPACKTT 407  
DB 363 EIRFIIDNEPSIDSTFKYKFDPSIMGNLEKFNHFNPSRSEVQILKGLNLKVKVSO 422  
QY 408 ALVAGSGSKSTWGLLEKRWYLPSSGRILLDGLGELGQYVNWKLRSRLRVQOEPVLP 467  
DB 423 ALVNSGCGKSTTVOLMORLYDPLGCVVSDIGODIRTINRYLRHIIIGVVSQEVLP 482  
QY 468 IFQNIANGPFMDQORLPREKQKELVQKACKASNGDVFINELPNGYETEVEGERAGAL 527  
DB 483 IAEINIRYGRD-----VTMDETEKAVKEANAYDFIMKLPHQFDLTVGERGAQLSG 534  
QY 528 RORITATARSTISDPKILLDEATSAIDPKAEKYVQOELNRYNSKDRTTLVIAHKLATV 587  
DB 535 KORIALARALVRNPKILLDEATSAIDTSEAVQOALDKAREGRTTIVIAHRLSTYRNA 594  
QY 588 GNIAVISQKIVBQTHHELIEFGCHYAALVRAQDLGADEQ-----QEHKTLREKAA 640  
DB 595 DVTAGDGGVIVEQGNHDELMREKGIYFKLVMTQTRGNEIEPGNNAYGSDSDTASE 654  
QY 641 REAAGERPALERTHTTATSQAGDLEKRVKPVGTGLYSLKCI-----LIMFYEOKNL 692  
DB 655 EES--KSLIRRSIYRSVHRKQDERR-----LSMKEAVDEVDPLVSEFWRLNLS 704  
QY 693 YWCFLI-SITVILICAATPPGALLFSLTLVFEELSG--HAAQERADPILMFVVAALGN 749  
DB 705 ENPILLVGLCAVINCCIOPVFAIVFSRLVGVFSRDDHETKRONCNLSLFFLVWGLIS 764  
QY 750 LVGYFTIGWTCNVISQVWTHRYQAAAFQRLVDQDELLDIDPEQISGALTSALPTQL 809  
DB 765 FVYTFYGGFTFGKAGBILTKRVYMYFKSMLRQDISWFDHDKNTGSLTTHLASDASS 824  
QY 810 ELISANFLIYVVGQHRLEQCSTTSLMMETPGCGVCMCTSTPAFCW----- 855  
DB 825 GAMGARLAV-----VTQNVANLGTGVIL-----SLVYGWQLTLLVVIPLIV 867  
QY 856 -----LPQNSRDEARSRLKCLCKCKWACKRSRSDPDRLLIFDSRRPCSPVLGH 906  
DB 868 LGGIEMKLLSGOALKDKKQLEISKIATEA-----IENFRTIVSLTREOK 913  
QY 907 VEOGLAKITQ-----SFWFGRCFGFHLISQSMELAIAGFCIAVDNWLQVSTTQLN 957  
DB 914 FETMYAOSLQVYRNAMKKAHVGTFTSF--TQAMMYFYAACFRFGAYLVAQOOLMTFEN 971  
QY 958 FISSWAFCLPVQAAAQYLAISTFTKARSANYILWL-----RTLKPTIRET 1005  
DB 972 VMLVFSVAVVGEAMAGNCTSEFAPDYAKAKVASHIIRIIEKTPEIDSYSTVEGLKPT 1031  
QY 1006 EENKKGPVGGCPVDLEDIEFRYRORDSARVLRGVSMITPEQPFVAVYVAGSGCKSLIA 1065  
DB 1032 N-----VKRNGVQFNYPTRPNIPVLQGLSLBYRKGOTLALVYVGGSGCKSTVQ 1079  
QY 1066 LSREFYDPTSGRISFAHENAEMSPRLYRGHMSLVQOQPTLYQGSVRENTLA--LEAEL 1123  
DB 1080 LLERFYDPMAGSVFLDGRKEIKQLNQMVRHLGIYVQSPELILFDCSIAENTAYGDNRAVS 1139  
QY 1124 SEELCCGRPLPAPMLWILSSLYOKALKRLAORQHQFSGGQORQRIATARALIRNPKILL 1183  
DB 1140 HEEIVRAAKEAN-IHQFIDSLPKYNTVR-GDKGTQLSGGQKORATARALVRPHILL 1197  
QY 1184 DEATSALDQSERLVOAALDEASTRTTIAVAHRLSTIRNVDVIFVFANGRIAGTGTAE 1243  
DB 1198 DEATSALDTESEKVVQOALDKAREGTCIVIAHRLSTIONADLIWVIENCKVKEHGT 1257  
QY 1244 LQRLGRYVEMCLA 1257







DB 178 VSKINDGIGDKIGMFQSIATFLAFLVGFISGWKLTILVAVSPLIGLSSAMWAKVLTS 237  
QY 222 QEARITLYSRANVAEVLSSITRTAFYAOQKMKVEKVDYFLOAOHQEGKKSPNNGVL 281  
DB 238 FTNELONAYAKAGAVEVLAIRTVIAGGQNKELERKNLEKKNVGIKKAVTANIS 297  
QY 282 FSTYFCIYAAIALAFKFMRYONGEADVKVFTVVLVSVTLAATSISMLAPSGSVYQ 341  
DB 298 TGIAYLIVYATALAFYGTSLVSNEYSS-VGOVLTVFTSILGFTSGHTAPNIEVFAN 356  
QY 342 RRIEGSELFSTIDKPTOLDPLDSQKQEGCIGQIEIQNALFAYPSAQVLRDFNLTI 401  
DB 357 ARGAAETFKIIDNEPSDFSOGHKPDSVMGNLEFKNVHFSYPSRGIGIKLGLNLKV 416  
QY 402 PAGKTTALVAGSGSGKTMVGLLRYLPSGRILLDGLGQYNNKWLRSRILRVOQEP 461  
DB 417 QSGQVALVKGSGGKSTTVOLLQRLYDTEGVSDIGQDRTINRYRLREIIGVSGEP 476  
QY 462 VLRGTTFQNTANGFMDQDLPREK-QMELVOKACKASNGDVFINELPNGETVEGGERA 520  
DB 477 VLFATTIAENIRYG-----RENVMTDEIEKAVKEANAYDFIMKLPKFDTLVGBRG 527  
QY 521 CALSGGQRIATARSIIISDPKILLDEATSDALPRAEKVVOEALNRVSKDRTTIVIAHK 580  
DB 528 AOLSGGOKRIATARALVRNPKILLDEATSDALPRAEKVVOEALNRVSKDRTTIVIAHR 587  
QY 581 LATYKSAGNIAVISQKIVGQTHHELIFEGCHYAALVBAQ-----DLGADEQOEHEKT 634  
DB 598 LSTVRNADVINGFGGVIVGQNHBELEKGIYCELVNMOTRGNEVELGSEADGSDOT 647  
QY 635 LHEKARPAACERPALERTWTATSOAGLEKR-----KVPVGTGLYSLLKILIM 685  
DB 648 TASELTSEEF-KSPSVKSTCRSICGSDQERRVSVKEAQDEVDPLVSF-WGILKLNIT-- 703  
QY 686 FVEOKNLYWCLP-LSTTVILICATFPQALLFSRLITVF--ELSGHAAQERADFIYLMF 742  
DB 704 -----TEWPLYVGLVCAVINGMPVFSIVFSIGIIVTRDDPKTKQKCNLFSLFF 757  
QY 743 FVALGNLVGFTYGTWCNIVSOVTHRYCAAMFORVLDDIETLDDIPFQISGALTQSL 802  
DB 758 LVMGICFVTFYFGFTGAGEILKRLRYWFKSMLQDLSWEDDHNSGTGALTFLA 817  
QY 803 ALPTQLBLISANFLIIVVQHRLEOCSTSLWMTGPGGVCWCTSTPAFGW----- 855  
DB 818 SDAANVKGAMSS-----RLAGITONVANLGTG-----IIISLVYGMQLTLLV 860  
QY 856 -----LPQNSRDEARSKLGLCKRCWKACKRSYRDPDLIFDSRRPC 899  
DB 861 VIAPLILSGMMENKVLGGQALKDKKELEVSGKIATEA-----IENFTVV 906  
QY 900 SPTVLGHVEOGLAKIIQ-----SEWFGRCFGPHLSOSMEFLAIALGFCIAVDNWL- 949  
DB 907 SLTREKFNENYAGSLQIPYRNALKKRAHVEGITFSF--TQAMMYFSYAA--CFREGAYLV 962  
QY 950 --QVSTTQNLNIISSWAFCLPVQAAQYLAYSTFTSKARSAANYILNLTUKPTI----- 1002  
DB 963 AHQIMTFE-NYMLVFSVAVFGAIAAGNASSFADPYAKAVSASHIRINEKIPSIDSVST 1021  
QY 1003 RETEENKKGPVGGCPVLEDIEFRYQRDSARVLRGVGMTIEPQGFVAYVAGSCGKST 1062  
DB 1022 RGLKPNWLEG-----NVKFNENVFNPTRPDIPLVQLGSLVKKGOTLALVGGSCGKST 1076  
QY 1063 LIALSERVYDPTSGRISFAHENAEMSPRLYRGHMSLVQOETPLXOGSVRENVTLALAE 1122  
DB 1077 VVOLLERYDPMAGTVFLDGSKEIKQLNVQWLRALHIVSQEILFDSCIAENIAYGDSNR 1136  
QY 1123 -LSEELCOGRUPARPMILWSSLYOKALKPLAQRQKQSGQRQRIATARALNPKLL 1181  
DB 1137 VVSQDIETRAAKEANIHQFIESLPDYNTRV-GDKGTQSGGQKQRIATARALVROPHL 1195  
QY 1182 LIDEATSALDQSERLQVQALIDEASTSRTIAVAHRLSIRNVWDVVFVANGRIETGTH 1241  
DB 1196 LIDEATSALDSEKVVQVQALDKAREGRTCTIVIAHRLSTIQNALDILVVIQNGKVEHGT 1255

QY 1242 AELQRLGRGVYEMCLA 1257

DB 1256 QOLLAQKGIYESWQA 1271

## RESULT 7

MDRI\_CRIGR  
ID MDRI\_CRIGR STANDARD; PRT: 1276 AA.  
AC P21448;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Multidrug resistance protein 1 (P-glycoprotein 1).  
GN ABCB1 OR PGY1 OR PGP1  
OS Cricetulus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetulus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92135896; PubMed=1685679;  
RA Endicott J.A., Sarangi F., Ling V.;  
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein  
RT gene family";  
RL DNA Seq. 2:89-101(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91154265; PubMed=1671863;  
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;  
RT "Full length and alternatively spliced pgp1 transcripts in multidrug-  
RT resistant Chinese hamster lung cells";  
RL J. Biol. Chem. 266:4545-4555(1991).  
RN [3]  
RP SEQUENCE OF 706-1276 FROM N.A.  
RX MEDLINE=88122132; PubMed=2893255;  
RA Endicott J.A., Juranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,  
Ling V.;  
RT "Simultaneous expression of two P-glycoprotein genes in  
RT drug-sensitive Chinese hamster ovary cells";  
RL Mol. Cell Biol. 7:4075-4081(1987).  
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
CC -!- DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT  
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY  
CC CANNOT.  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DB EMBL; M60040; AAA68883.1; -  
DB EMBL; M59253; AAA37004.1; -  
DB EMBL; M17897; AAA37006.1; -  
DB PIR; A38696; DVHY1C.  
DB InterPro; IPR003593; AAA.  
DB InterPro; IPR001140; ABC\_transporter\_tmam.  
DB InterPro; IPR003439; ABC\_transporter.  
DB InterPro; IPR001687; ATP\_GTP\_A.  
DB Pfam; PF00664; ABC\_membrane; 2.  
DB Pfam; PF00005; ABC\_tran; 2.  
DB SMART; SM00382; AAA; 2.  
DB PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
KW Multigene family.  
FT DOMAIN 1 50 CYTOPLASMIC (POTENTIAL).

```

FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 344 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 754 774 POTENTIAL.
FT TRANSMEM 830 860 POTENTIAL.
FT TRANSMEM 831 871 POTENTIAL.
FT TRANSMEM 934 954 POTENTIAL.
FT TRANSMEM 971 991 POTENTIAL.
FT TRANSMEM 992 1276 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1276 140925 GA -> AP (IN REF. 2).
FT NP_BIND 424 431 ATP (POTENTIAL).
FT NP_BIND 1067 1074 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1276
FT CONFLICT 338 339
SQ SEQUENCE 1276 AA; 140925 MW; 44F3F92A186B4DEF CRC64;

Query Match 25.0%; Score 1612; DB 1; Length 1276;
Best Local Similarity 32.1%; Pred. No. 1.2e-96;
Matches 413; Conservative 251; Mismatches 534; Indels 88; Gaps 18;

QY 24 EKETEGISDYLRIFRYADKYDWTUNLVNICALICAGSGASPLMSLIIFSGSTNKF----- 77
DB 25 EKKEKPPVSVFTMPRYAGWLDRLVMTGLAAIHGVALPLMLVFGDMTDSFASVGN 84
QY 78 -----NNYASGGSGPEAKADVDFHVLVFFVLFIFGKVL--TVYSAAITISAIRTRTL 130
DB 85 PTNATNATOVNASDIFGKLEEMTYAYVYTGAGVLVAYIQVSPWCLAAAGROIHK 144
QY 131 RYVELECTROEWEHDFKOSNCAIATVXTNGNRIQTGIAEKLVTFTVQALSWFESFVVA 190
DB 145 RQKFFHAINQHEIGHDFVNDVGLNTRTDDVSKINGIGDKGMFOFQATMFFGGFIIG 204
QY 191 LASQWKLALITWSPITAFVLTGICIAIDAAQEARITRYSRAVLAEEVLSSTIRTVAF 250
DB 205 FTGKWLTLVILAIAPVLSGLSAGIWKILSSFTDKELQAYAKAGAAVEVLAIRTVIAP 264
QY 251 YAKKMKERYDYFLOQAQHEGKKSPNNGVLFSTFEYCIYAAIALAFWKFGRWYQGEVA 310
DB 265 GGOKKELERYNNLEAKRLGKAITANISMGAAFLIIVASYALAFWGTSLVISKEYS 324
QY 311 DVGKVFVTVLTLATSTLSMLAPSGSVVYQRRIFGSELSIIDKPTQLDPLDPGSKOPE 370
DB 325 -IGQVLTVFVAVLIGAFSIGQASPNIEAFANAGAAVEIENIDNPSIDSFESKNGYKPD 383
QY 371 GCLQTEIONLAFAPSPRSQVLDNFNLTPAGKTTALVGASGSKSTNKGVLKRYLP 430
DB 384 NTKGNLEFNHIFSPSRKRDQVILKGLKLVQSGQTVLVNLSGCGKSTTVQLQLRLYDP 443
QY 431 SSGRILLDGLGOYNVVKWLSRIRLVQOEPVLFGRGTFIQNIANGFMDEORDLPREK-QM 489
DB 444 TEGVWSIDGQDRTINVRVLEIGVVSQEPVLFATTAENIRY-----RENVTM 494
QY 490 ELVOKACKASNGDVFINELPNQYETVEGERAGALSGGORORAIARSIIIDPKILLDEA 549
DB 495 DETEKAVKEANAVDFIMLPKPHKFDLVGERGAQLSGSGKORAIARALVRNPKILLDEA 554
QY 550 TSALDPKAEKVQOEAALNRVSRDRTLVIAHKLATVKVAGNIIVISQKIVEQGTHELLE 609
DB 555 TSALDTESEAVVOAALDKAREGRTIVIAHRLSTVRNADIAGDGVGVYQGNHEELMR 614
QY 610 FCGHYAALVRAODLQAD-----EQDHEKTLHEKAAAREAGGERPALERTHTTATSOAG 662
DB 615 EKGIFPKLVMTQAGNEIELGNEVESKNEINDLMDSKOSAS---SLIRRRSTRSIRG 671
QY 663 --DLEKRPVGTGLYSLLKCLIMFYEQKNLYW-CFLLTITVLTCAATFPQOALLFSR 719
DB 672 PHQDRKLTKEALDEDVPPISFWRILKLSNSENPEYVVGIFCAIVNGALQPAFSIIFSK 731

```

```

QY 720 LLTVFELSCHAQERAD--FYILFFVVALGNLVGYFTIGWTCNVISOVVTHRYQAAAFQ 777
DB 732 VVGVTNRNTDDETKRHSNLSLFLILGVSIFITFFLQGGTFGKAGELTKRLRYVVK 791
QY 778 RVLDDQIELDIPQISGALTSQLS-----ALPTQLQELISANFLIYVWG 823
DB 792 SMLRQDVSNFDPKNTTGALTTRLANDACOVKGATGARLAVITQNIANICGTIISLIYG 851
QY 824 QHRLQOCSTTSLMMETGPGCCVACTSTPAPGWLPONSSRDEARSRKGLKCRKCKWACKRS 883
DB 852 -----WQLTLLLLAIYP--IIALAGVVMKMLSQALKKKELEGGCKATEA----- 897
QY 884 SYRDPRLIPDSRRPCSPVTLGHVPOGLAKITQ-----SPWFGRCFGFHLQSQMEF 934
DB 898 -----IENFTVVSILTREQKFENMYAQSILQIPYRNALKKAHVFGITFSF--TOANMY 947
QY 935 LAIALGFCIAVDNMLQVSTTQNLFTSSSWAFCLPVQAAQYLAYSTFTFKARSAAVILW 994
DB 948 FSYAACFRGAYLVARELMTFENVLVFSIVFGANAVGOVSSPAPDYAKAKVSASHIM 1007
QY 995 LRTLKPTIRETEENKKGPVGGCPVDLEIDIEPRKQORDSARVLGVSWMTIEPGQFVAVG 1054
DB 1008 IIEKVFSIDSYSTGGLKPNTEGNKVFNEYVFNPTRPDIPVLOGLNLVKKGGOTLALVG 1067
QY 1055 ASGCGSTLIALSERPYDPTSGRISSFAHENAEMSPRLYRGHMSLVQOEPFLYOGSVREN 1114
DB 1068 SSGCGKSTVQLLEFRYDPMAGTVFLDGKEVNLVQVRAHLGLIVSQEPILEPCSAEN 1127
QY 1115 VTIALBAE-LSERLCQGRLPARPMLWLSLSSVQALKRLAAQRMQFSGGQRORIALARA 1173
DB 1128 IAYGDSRVVSQDEIERAAKEANIHOFTESLPQKYNTRV-GDKGTQLSGGQKORIALARA 1186
QY 1174 LIENPKLLLDLDEATSDALDQSERLVQAAALDEASTSRTTIAVAHRLSTIRNVDFVIFVANG 1233
DB 1187 LVQPHILLDDEATSDALDTESEKVVQEAALDKAREGRTCVIAHRLSTIONADLIVVQNG 1246
QY 1234 RIATGTHAEQLRLQRYEYEMCLAOS 1259
DB 1247 VKKEGTHOOLLAQKGIYFESMVSVQA 1272

RESULT 8
MDR2_MOUSE STANDARD; PRT; 1276 AA.
AC P21440;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 2 (p-glycoprotein 2).
GN PGY2 OR PGY-2 OR MDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88302195; PubMed=3405218;
RA Gros P., Raymond M., Bell J., Housman D.;
RT "Cloning and characterization of a second member of the mouse mdr
RL gene family.";
RL Mol. Cell. Biol. 8:2770-2778(1988).
RN [2]
RP SEQUENCE OF 1-23 FROM N.A.
RC STRAIN=BALB/C;
RA Kirschner L.S., Horwitz S.B.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS
CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.

```







```
QY 988 AANYILMLRTKPTIREB-----ENKKPGVGGCPVOLEDIEFRYRORDSARVLRGVSM 1042
Db 1005 SAHLFSLFRQPLDISYSGELWDPKEG-----SVTFNEVVFTRAMPVQLQSL 1059
QY 1043 TIEBCQVAVVAGSCGKSTLIALSERYPDTSGRISFAHENIAEMSPRLRGHMSLVQ 1102
Db 1060 EVKAGOTLALVGGCGKSTVOLLERYPDMAGTLLDGOBAKKNLQMLRAQLGIVSQ 1119
QY 1103 EPTLYQSGVRENTVLADEL--SEELCQGRLEP--RPMWLSSLYQKALKRLAAQRM 1158
Db 1120 EPVLFDCSIAENIAYGDSRVVSODEIVRAKAAHTPIETLPQKY-----KTRVGDKGT 1175
QY 1159 QFSGGQRATATARLIRPKLLDDEATSLQTSERLVQALDEASTRTTIAVAHRL 1218
Db 1176 QLSGGQRATATARLIRPKLLDDEATSLQTSERLVQALDEASTRTTIAVAHRL 1235
QY 1219 STIRNDVIVFANGRIATGTHAELQRLGRYVENCCLAQS 1259
Db 1236 STIQNADLVVIONGKVEHGHOOQLLAQKGIYFSMVNIQA 1276

RESULT 11
MDR1_RAT
ID MDR1_RAT STANDARD: PRT: 1277 AA.
AC P43245;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 1 (p-glycoprotein 1).
GN ABCB1 OR PGY1 OR MDR1 OR MDR1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RW [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92039081; PubMed=1682220;
RA Silverman J.A., Raunio H., Gant T.W., Thorgirsson S.S.;
RT "Cloning and characterization of a member of the rat multidrug
resistance (mdr) gene family.";
RL Gene 106:1229-236(1991).
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
LINKER DOMAIN.
CC -!- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
RELATED BUT DISTINCT CELLULAR GENES.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M81855; ; NOT ANNOTATED_CDS.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC transporter_tmern.
DR InterPro: IPR003439; ABC transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; glycoprotein; Transmembrane; Transport; Repeat;
FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 71 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
```

```
FT TRANSMEM 192 219 POTENTIAL.
FT TRANSMEM 236 316 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT DOMAIN 348 709 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 710 730 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 832 852 POTENTIAL.
FT TRANSMEM 854 874 POTENTIAL.
FT TRANSMEM 937 957 POTENTIAL.
FT TRANSMEM 968 988 POTENTIAL.
FT DOMAIN 989 1277 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (BY SIMILARITY).
FT NP_BIND 1070 1077 ATP (POTENTIAL).
FT REPEAT 636 1277
FT CARBOHYD 91 91 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1277 AA; 141386 MW; 8AFDD619D2934C1 CRC64;

Query Match 24.6%; Score 1590; DB 1; Length 1277;
Best Local Similarity 31.7%; Pred. No. 3,3e-95;
Matches 424; Conservative 254; Mismatches 510; Indels 148; Gaps 30;

QY 3 DESEKPRNODGSESSHPPEKTECSI DYLRIRFVADKYDWTLVNVLALCAIGSGAS 62
Db 4 EELGNRADKNFSGMKKSKKEKPAVCIF-GMFRVADWLDKLMALGTAAIIHGTL 62
QY 63 LPLMSIIIGSTNKF-----NYNNSGDG-SPEAFKADVDFVLFVLF 106
Db 63 LPLMLVFGYMTDSFTPSRDPHSDRAITNQSEINSTHTVSDTSLEEDMAYAYITGIGA 122
QY 107 GFRTVTVYSTAAITSAITRTLRVFLVLECTLRQVNHFPKQNGAIXATVTTNGRRT 166
Db 123 GVLIVAYTQVSLWCLAGROIHKIRQKFFHAIIMNOEIGCFDVPDAGELNRLTDDVSKIN 182
QY 167 TGIAEKLVFTVQALSMFSAFVVALASOMKALITMSVIPAFILVTGICIAIDAAQEAR 226
Db 183 DGIGDKLGMFQSITTSAGFIIGTSGWKLLTVLAVSPILGLSSAMAKVITSTNKE 242
QY 227 TRIYSRAVLAEEVLSSIRTVHAFYAQKKMKVEKYDFVFLQQAHOEKKKSPNGVLFST 286
Db 243 LQAYAKAGAAVEEYLAARTVIAFGGQKKELEKRYNKNLEAKRVGINKKAITANISIGI 302
QY 287 FCYIAALAFWKGFRMYONGEADVGVKVTYVLSVTLAATSIISMLAPSGVYVQRIRFG 346
Db 303 LLVYASALAFWYGTSLVLSNEYS-IGQVLTVFFSILLGTFESIGHLAPNTEAFANARGA 361
QY 347 SELFIIIDKPTQLDLPDPSKGQPEGCLQIETIQNLAFAYPSRPSAQVLRDNLTIPAGKT 406
Db 362 YEIFKIIDNEPSIDSFSTKCHKPDSIMGNLEFKNVYFNYPSPSEVKILKGLNLKVKSGOT 421
QY 407 TALVCASGSGKSTWVGLLEWYLPSSGRILLDGLGQYNNVKNLRSIRLVQOEPVLRPG 466
Db 422 VALVNSCGKSTTVQLQLXDPIEGEVSIDGDIRTINRYLREITIGVVGSEPLFAT 481
QY 467 TLFONIANGFMDQORDLPREK-QMELVOKACKASNGDVFINELPVGVEYEVGERACALGS 525
Db 482 TIAENIRVG-----RENVTHDEIEKAVKEANAYDFIMKLPHEFDTLVGERGQLSG 532
QY 526 GORQBIATRSIISDPKILLLOEASALDPRAKVVQALNRVSKDRTLVIAHKLATVK 585
Db 533 GORQBIATRALVRNPKILLDEATSEAVYQALDKAREGHTTILVIAHRLSVIR 592
QY 586 SAGNTAVTSQKIVVEQCTHHELFQCHYALVRAQ-----DLGADP- 627
Db 593 NADVTAGDFGGVIVVEQGNHEELMKEGYIFKLVMTPTQRTGNEIEPGNNNAVESQSDTGASL 652
QY 628 -QOEHEKTLHEKAAAREAGERPALEHTTATTSQAGDLEKRVKVPVGLVSLKLCILINF 686
Db 653 TSEESKPLIRRSIRKSIHRRODQER---RUSSKEDVDE-DVP-----MWSF 695
QY 687 YEQRNL---YWCFL-LSTITVLICATTPFGQALLFSRLTLVFEF-SCHAAQER-ADFYIL 740
```

```
Db 696 WQILKLNISEPVLVGVLCVINGCQVFAIVFSGVGVSRDDHETKQRCNLFSL 755
Qy 741 MFEVVALGNLVGTIGWTCNVSQVVRHRYOAAAMFORVLDLDDLELDIDPEQISGALTSO 800
Db 756 LFLVNGMISFVTFQGFTEGKAGELIKRLRYAVFVSMRLRODISWFDHKNFTTSLVTR 815
Qy 801 LSAIPQLOELISANFLIIVVQHRLEOCSTSLWMTGGCVCTSTPAGW----- 855
Db 816 LASDASNVKGAWSRLAV-----VTQNVANLFG--IILSLVLTVMQTL 860
Qy 856 -----IPQNSSRDEARSKLGLCKKCKWACKRSYRDPRLIFDSRR 897
Db 861 LVIIPLVLGGIIEKMLLSGQALKDKKELEISGKIATEA-----TENFT 906
Qy 898 PCSPTVLGHVROGLAKTIQ-----SFMFGRCFGFHLISQSMEFLAIALGFCIAVDNM 948
Db 907 VVSLTREQKFTMTAQSLQIPYRNALKKAHVGITFAP--TQAMIVFSTAA--CFREGAY 962
Qy 949 LQVSTQLOLNFSSWAPCLPV---OAAQYLAYSFTSKARSANYILNLRLTKPTI--R 1003
Db 963 L-VARELMTFENVMLVPSAVVFGAAGNTSSFADYAKAKVSAASHIIGIEKIPEDSY 1021
Qy 1004 ETEENKKGKPVGCPVDELDIEPRYORDSARVLRGYSMTLEQCFVAYVAGSGCKSPL 1063
Db 1022 STEGLKPNMLEGN--VYFNCVKNYPTNPVLPGLSFEVKKQTLRLVGSNGCKSKSV 1079
Qy 1064 IALSERFDPTSGRISPAHENIAENSPLRYGHNSLVQOEPFLYQGVRENVTLEABL 1123
Db 1080 VQLLEKRFNPMAGTVFLDGKEIKQLNVQCVRA--LGIVSQEPFLDCSIAENIAYGDSNRV 1138
Qy 1124 --SEELCOGRUPARPMLWILSSLYOKALKRLAAQRMQFGSGGQRIATARALINPKLL 1181
Db 1139 VSHERIVRAAREAN--IHQFIDSLPERKYNTRV--GDKGTQLSGGQRIATARALVROPHIL 1196
Qy 1182 LDEATSAIDTQSERLVOALDEASTRTTIAVAHRLSTIRNVDFVFPANGRIATGVR 1241
Db 1197 LDEATSAIDTQSERLVOALDEASTRTTIAVAHRLSTIRNVDFVFPANGRIATGVR 1256
Qy 1242 AEOLRLGRYYEMCLA 1257
Db 1257 QQLLAQKGIYFSMVOA 1272

RESULT 12
MDRL_CAEEL STANDARD: PRT: 1321 AA.
AC P34712;
Dt 01-FEB-1994 (Rel. 28, Created)
Dt 01-FEB-1994 (Rel. 28, Last sequence update)
Dt 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein A).
GN PGP-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=93085750; PubMed=1360540;
RA Lincke C.R., The I., van Groenigen M., Borst P.;
RT "The P-glycoprotein gene family of Caenorhabditis elegans. Cloning
RT and characterization of genomic and complementary DNA sequences.";
RL J. Mol. Biol. 228:701-711(1992).
RN [2]
RP TISSUE SPECIFICITY.
RC STRAIN-BRISTOL N2;
RX MEDLINE=93223702; PubMed=8096815;
RA Lincke C.R., Broeks A., The I., Plasterk H.A., Borst P.;
RT "The expression of two P-glycoprotein (pgp) genes in transgenic
RT Caenorhabditis elegans is confined to intestinal cells.";
RL EMBO J. 12:1615-1620(1993).
```

```
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: INTESTINAL CELLS.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X65054; CAA46190.1; -.
CC PIR: S27337; S27337.
CC InterPro: IPR003593; AAA.
CC InterPro: IPR001140; ABC_transporter_tmem.
CC InterPro: IPR001439; ABC_transportr.
CC InterPro: IPR001687; ATP_Gtp_A.
CC Pfam: PF00664; ABC_membrane; 2.
CC Pfam: PF00005; ABC_tran; 2.
CC SMART: SM00382; AAA; 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 2.
CC ATP-binding: Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
KW DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 321 341 POTENTIAL.
FT TRANSMEM 350 370 POTENTIAL.
FT DOMAIN 371 753 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 754 774 POTENTIAL.
FT TRANSMEM 798 818 POTENTIAL.
FT TRANSMEM 874 894 POTENTIAL.
FT TRANSMEM 895 915 POTENTIAL.
FT TRANSMEM 978 998 POTENTIAL.
FT TRANSMEM 1017 1037 POTENTIAL.
FT DOMAIN 1038 1321 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 451 458 ATP (POTENTIAL).
FT NP_BIND 1112 1119 ATP (POTENTIAL).
SQ SEQUENCE 1321 AA; 145074 MW; 5130AFA3B31ABFA9 CRC64;
```

```
Query Match 24.0%; Score 1551; DB 1; Length 1321;
Best Local Similarity 30.9%; Pred. No. 1.2e-92;
Matches 411; Conservative 245; Mismatches 500; Indels 172; Gaps 30;

Qy 36 RIFRYADKYDWTNLNVALICAGSGASLPLMSIIFGSPFNKF----- 77
Db 64 QLYRYTTTLEKLLFGTLVAVITGAGLPLMSIQGKVSQAFINEQIVINNGSTFLPTG 123
Qy 78 NNYSGGSGPEAFKADVDHFLVFLFYFLFGKFLVTVYSTAATITSAITRTTLRRVFLFC 137
Db 124 QNYTKTD-----FEHDMVNVVSYAAATVGMMAAGQIIVTCYLVABQMNRLRREFVKS 178
Qy 138 TLROEVWHFKQSGNGLATVTTNGNRIOTGIAEKLVTQALSMFSAFVALASQMKL 197
Db 179 ILRQESWEDFNHSGTLATKLEDFNLERKVGEGDKIGMAFOYLSQFITGFIATFTHSMQL 238
Qy 198 ALITMSVTPAFLVTCGICIAIDANQEA-RITRIYSRAAVLAEEVLSSTIRTVHAFYAKKM 256
Db 239 TLVLVATVP--IQALCGFAIAKSNSTFAIRETLRYAKAGKVVEETISSIRIVVSLNLRVE 297
Qy 257 VEKVDVFLQQAQHEGKKKSPNNGCVLFSTFYFCIYAAIALAFWKGFYQNGEVADYGVKVF 316
Db 298 LERYSTAVEEAKAGYVLKGLFLGISFGAMQASNFISFALAFYIGVGVHDSGL-NFGDML 356
Qy 317 TVVLSVTLATFSISMLAPSGSVVYQRIFFSGSELSIDKPTQLDPLDPGSKQPECCGCGI 376
Db 357 TTFSSVMNGSMALGAPQLAVLGTAGGAAGSIYEVLDKRPVIDSSSKAGKDKMTKGD 416
```

```
QY 377 ETQNLAFAYPSRPAQVLRDNLFTIPAGKTTALVAGSGGKSTWVGLRWLWLPSSGRIL 436
DB 417 TVRHFYTPRPOVPIRGMRVNRAGQVVALVSGSGGKSTIIISLLRYDVLKQKIT 476
QY 437 LQGLGQYNNKWLRSRIRIVQDEPVLFRGTIPONTANGFMQDEQDLPRKQKQVQAK 496
DB 477 IGDVDRDINLEFURKVVAVSDEPALFNCITTEENISLG---KEGITREMW---AAC 528
QY 497 KASNGDVEINELPNGETYECERAGALSGGQRIARISIIISDKILLLEATSAIDPK 556
DB 529 KMANAEKIKTLPNGYNTLVGDRGTQLSGGQRIARALVRNPKILLLEATSAIDAE 588
QY 557 AKVVQVQALNRVSDRTTLTAHKLATVKSAGNIASVISOGRKIVEGTGTHLIEFGCHAA 616
DB 589 SEGIVQQAALKAAGRTTIIAHLSTIRNADLIISCKNGOVVEGDHRLMAQOGLYD 648
QY 617 LVRAQ-----DLGADEQOHEKTL-----HEKAAREAGERPALERTHTTATSQAG- 663
DB 649 LVTADTFDAVSAEKGFSRENSVARQTSHEGJSRQNSMDDIMNRVRSSTIGSTNG 708
QY 664 --LERKVPVTLGYSLIKCIL-----INFYEQRNLYWCFL-LSTIVLCAAA 708
DB 709 PVIDEKEERIKDALKSLKQLENNNAQTNLFELIYHARPHALSIFGMSTAT--IGGF 766
QY 709 TFPQALLFSLRVFELSRAAQ--ERADFYILFFVVALGNLVGYFTIGWTCNVISQV 766
DB 767 IYPTVSVPFTSFNMVE--AGNPADLSQGHFALMFLVLAQAQGCSCFLMTFFMGIASES 824
QY 767 VTHRYQAAMFORVLDQDIELDIPQISGALTQSALPTLOELISANFLIYIVVQQR 826
DB 825 LTRDLRNKLFRNLVLSQHIGFSDPQNASCKISTRLATDVNLRATIDRFSTVI----- 878
QY 827 LEOCSTTSLWMETGPGCVCTSTPAFGW-----LPQNSSRDEARSKL-GKLCRK 876
DB 879 -----TTLVSMVAGIGLAF-----YGNOMALLIIILPIVAGQYLRGRFTGKNVKS 927
QY 877 CHACKRSYRDPDLIFDS-----RRPCSTVLCHVDEQGLAKIIQSFV 919
DB 928 A-----SEFADSGIAIEATENVRTQALAREDTFYENPCEKLDIPHEAKIEAFIOGLS 982
QY 920 FQRCGFHLSQSMFL-----AIALGFCIAVDN-----WLOVSTTOLNFISS 961
DB 983 YG-----CASSVYLINTCAYRMGLALIITPPTQPMRVLRVWYAITSTPLGFSATS 1036
QY 962 SWAFCLPVAAAQYAVSTSTKARSAANYIL-WLRTLKPTIRETEENKKGPPVGGCPVD 1020
DB 1037 -----YFPEYAKATFAGGIIIFGMLRKISKIDSLSLAGEKKLYG--KVI 1078
QY 1021 LEDIEFRYRDSARVLRGVSMTEPQGFVAVYGASGGGKSTLIALSERFYDPTSGRISF 1080
DB 1079 FKNVRFAYPERPEIEILKGLSFSVEPQTLALVGPSCGKSTVALLERFYDTLGGEIFI 1138
QY 1081 AHENIAEMSPRLYRGHMSLVQOQPTLYOGSVRENTVLALEAELESELCOGRLPARPMWI 1140
DB 1139 DGSSEIKTLNPHSTRSIAIVSQEPTLFDSCSIAENIIYGLDPS-SVTMAQVEEAR----- 1192
QY 1141 LSSLYQ-----KALKRLAQRGMQESGGQRIARIALRNPKILLLEATSAIDTOS 1194
DB 1193 LANIHNFNFAELPEGFYRVEDRGTQLSGGQRIARIALRNPKILLLEATSAIDTOS 1252
QY 1195 ERLVQAALDEASTRTTIAVHRLSRIRVNDVIFVPANGRTAETGTHAEQRLRGRIYEM 1254
DB 1253 EKVVOEALDRAREGRTCIVIAHRLNLTVMNADCIIVVNSNGTIEKGTHQTLQMSXGAYYKL 1312
QY 1255 CLAQSLOQ 1262
DB 1313 TOKOMTEK 1320
```

RESULT 13  
MDR4\_DROME

STANDARD; PRT: 1302 AA.

```
AC Q00449;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein homolog 49 (P-glycoprotein 49).
GN MDR49.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN NCBI_GeneID=7227;
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RC MEDLINE=91304385; PubMed=2072901;
RA Wu C.-T., Budding M., Griffen M.S., Croop J.M.;
RT "Isolation and characterization of Drosophila multidrug resistance
RT gene homologs."; Mol. Cell. Biol. 11:3940-3948(1991).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59076; AAA28679.1; -.
DR PIR: A41249; A41249.
DR FLYBase: FBgn0004512; Mdr49.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_ttfan; 2.
DR SMART: PS00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding: Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
KW DOMAIN 1 46 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 47 67
FT TRANSMEM 117 145
FT TRANSMEM 193 213
FT TRANSMEM 221 240
FT TRANSMEM 301 322
FT TRANSMEM 340 360
FT DOMAIN 361 734 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 735 756
FT TRANSMEM 780 801
FT TRANSMEM 856 876
FT TRANSMEM 878 897
FT TRANSMEM 960 980
FT TRANSMEM 997 1017
FT DOMAIN 1018 1302
FT NE_BIND 437 444
FT NE_BIND 1094 1101
FT REPEAT 1 670
FT REPEAT 671 1302
FT CARBOHYD 101 101
FT VARIANT 697 697
FT VARIANT 712 712
FT VARIANT 952 952
SQ SEQUENCE 1302 AA; 142724 MW; 4720C169455C28CD CRC64;
N-LINKED (GLCNAC... ) (POTENTIAL).
S -> N.
D -> N.
V -> I.
```

Query Match 23.4%; Score 1507.5; DB 1; Length 1302;  
Best Local Similarity 31.0%; Pred. No. 7.7e-90;  
Matches 426; Conservative 219; Mismatches 496; Indels 233; Gaps 37;



QY 28 EGSISD-----YLRIFRYADKYDWTLMVIALICAGSGASLPLMSILFGSFTNKFNNYN 81  
Db 19 EGSVDATRKYSYDLEFRYTRCERFLVLSLLVATAASAPFIPFMIYIGETSULLVORT 78  
QY 82 SGOG-SPRAF-----RADVDHFLVFWFLYFVIGFVFLVYVSPAAT 120  
Db 79 VGVGSSPAPALPFWGGGOQLFNASKENNAQIIDDATAFGSGSLVGSVAMFLITLAD 138  
QY 121 IS---AIRTTTLRVELCTLRQVWHFQKQNGAIATVXTNGNRITQIGIAKLVTV 177  
Db 139 LANRALNIDIRIKLFLEAMLRQDIANTDSSGNFASKMTEDLDKLEGIGKIVIV 198  
QY 178 QALSMFSAFVVALASQWKLITMSVPAIFLVYTGICIAIDAAQPARITRIYRAAVLA 237  
Db 199 FLMTFVIGIVSAFYQWKLITVLVSCVFFIAATSVVARLQSLAEKELKSYDAANV 258  
QY 238 EYVLSIRTVAFYQAQKMYKYDVPLOAHQHGKKSPPNGVLFSTEFYCIYAAIALAF 297  
Db 259 EEFVSGIRTVAFSGQEKEREGKULLIPAEITGRKKGLYSGMGNALSLIYLCMALAI 318  
QY 298 WRGFRM-YONGEVADVGKVP-----VVLVSYLAAYSISMLAPSGSVYQRIFGSELF 350  
Db 319 WYCVTLILDERDLDP--RVYTPAVIVLFAVINGAQNGLGFSHPHVEALAVATAAGOTLF 376  
QY 351 SIIDKPTOLDPSPGQPCGICQIETQNLAFAYPSPRPAQVLRDNFTIPAGKTALY 410  
Db 377 NIIDRPSQVDMDEKGRNPENTAGIRFEGIRYPARPQVEILKGLTVDVLPQOTVAFV 436  
QY 411 GASGCKSTMVGLLERNYLPSSGRILLDLGELGOYNVKKLSRIRLVQOEPVLFGRITFQ 470  
Db 437 GASGCKSTLIQLMQRYDEAGSVKLDGRDLRTLVNGLRSQIGVVGQEPVLFATFGE 496  
QY 471 NIANGPMDQRDLREKQMLVQKACKASNGDVFINELPNGYETEYGERAGALSGQOR 530  
Db 497 NIRYG-----RPSATQAD-IEKAAANACHDFITRLPXGYDTQVGEKGAISGQOKR 548  
QY 531 IAIARISIDPKILLDEATSLDPAKRVVOEALNRVSKDRITLVIAHKLATVKSAGNI 590  
Db 549 IAIARALVQPVQLLODEATSLDPTSEKRVQSALELASQPTLVVAHRLSTINADKI 608  
QY 591 AVTSQKIVQEQTHIELFGCHYAALV-----RAQDLG----- 624  
Db 609 VFLKGVAEQOHELMKERRGLYCELVSIORKEATEADGAVAGRPLOKSONLSOET 668  
QY 625 ----ADEQOEHEKTLHEKAAAREAGERPALERHTHTATSGADLEKRRKVPVGTGLYSLK 690  
Db 669 DODEDEDEPELOTSGRSDSGFRASRKR---RSORRRKKKKEKRVSKVSTOL- 724  
QY 681 CILIMEOKNLYWCF-LSTITVLICAATFPGQALLFSRLLTVFELSGHAAQERAD-FY 738  
Db 725 ----MKLNSPENRFIVGSGIASVMHGATFPLMGLFPGDFGILS-DGDDDDVVRAEVLK 777  
QY 739 ILMFFV-----VALGNLVGYFTIGWTCNVISQVTHRYQAMFORVLDQDIELDIPRO 793  
Db 778 ISMIFVIGLMAGLGNMLQYMF-----TTAGVKMTLRLKRAFGTIIGODIAYFDDENS 833  
QY 794 SGALTSQLSA-----LPTQLQELISANFLIYVVGQHRLEQCSSTSLWMEFGP 841  
Db 834 VGLCSRLASDCSNVOGATGARVGTMLQAV--ATLVGKVVG--FVFSMQOQLTLVLTLP 889  
QY 842 GGVNCTSPARGWLPQNSR-----DEARS-----BKLGKLC----- 874  
Db 890 ----LVCLSVYLEGRFLMKSKAKASIEASQVAVEATNIRVINGCLIERQVLDQYVQO 946  
QY 875 --RKCWACKRSYRDPDELFDSPRRPCSPVLVGHVEQGLAKITQSFWRGRCFGLHSQM 932  
Db 947 IDRVVDACRR-----KVPF-----RGLV-----FALGQA 971  
QY 933 EFIALIAL-----GFCIAVD--NMLQVSTTQLNFISSSWATCLPVQAAAQYLAYSTFTKA 985  
Db 972 PFLAYGISMYGGILVAERNYEDIIKVAEALIFGSMML-----GQALAYAPNVNDA 1024  
QY 986 RSAANYI--LWLTRLKPTIRETBENKKGPVGGC-----PVDULEIEFRYRQDSARVLR 1038

Db 1025 ILSAGRLMDLFKRT-----STQNPQSPYNTVKESEGDIVENVGYEYPTKRGTPILQ 1078  
QY 1039 GYSMTLEPGQVAYVYCASCKSTLIALSERFYDPTSGRISFAHENAEMSPRLYGHMS 1098  
Db 1079 GLNITIKKSTVALVPGSGSKSCVQLLRYDVPVSGVNLGSPVSTPEPLOTLSKLG 1138  
QY 1099 LYQOEPITLQSGVRENTLA-----LEAELSEELCOGELPARPMLWILSLVQ 1146  
Db 1139 LVSEPVLPDRTIAENIAYGNFRDDVSMQEIIEAAKSNHN-----FISALPQ 1188  
QY 1147 KALKRLAQRGMQFSGGQRIATARALINPKLILLDENTSALDQSERLYQAAALDEAS 1206  
Db 1189 GYDTRLG--KTSQSGQQRATARALVNRPKILLIDETSALDSEKVKVQALDEAR 1246  
QY 1207 TSRTTIAVHRLSTIRNVDFVFPANGRIETQTHAFLORLGRYYEMCLAQSL 1260  
Db 1247 SGRTCLTIAHRLTTVNRNADLCVLRGVVYVHGHDELMALNKIITANLYLMQV 1300  
RESULT 14  
ABILL\_RAT  
ID ABILL\_RAT STANDARD; PRT; 1321 AA.  
AC 070127;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)  
DE (Sister of P-glycoprotein).  
GN ABCB11 OR BSEP OR SGP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=101116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
RX MEDLINE=98212048; PubMed=9545351;  
RA Hofmann A.F., Meier P.J.;  
RT "The sister of P-glycoprotein represents the canalicular bile salt export pump of mammalian liver.";  
RL J. Biol. Chem. 273:10046-10050(1998).  
CC -!- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS INTO THE CANALICULUS OF HEPATOCYTES.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN SITU.  
CC -!- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES, EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN ATP BINDING CASSETTE (ABC) DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY, MDR SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
CC -----  
CC EMBL: U69487; AAC40084.1;  
CC InterPro: IPR003593; AAA;  
CC InterPro: IPR001140; ABC\_transporter\_unem.  
CC InterPro: IPR003439; ABC\_transporter.  
CC InterPro: IPR001687; ATP\_GTP-A.  
CC Pfam: PF00664; ABC\_membrane; 2.  
CC Pfam: PF00605; ABC\_tran; 2.  
CC SMART: SM00382; AAA; 2.  
CC PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
CC ATP-binding; Transmembrane; Transport.

```

FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 63 83 POTENTIAL.
FT DOMAIN 84 147 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 148 168 POTENTIAL.
FT DOMAIN 169 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 POTENTIAL.
FT DOMAIN 237 240 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 241 261 POTENTIAL.
FT DOMAIN 262 319 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 320 340 POTENTIAL.
FT DOMAIN 341 353 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 354 374 POTENTIAL.
FT DOMAIN 375 755 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 756 776 POTENTIAL.
FT DOMAIN 777 794 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 795 815 POTENTIAL.
FT DOMAIN 816 869 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 870 890 POTENTIAL.
FT DOMAIN 891 911 POTENTIAL.
FT TRANSMEM 912 979 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 980 1000 POTENTIAL.
FT TRANSMEM 1001 1011 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1012 1032 POTENTIAL.
FT TRANSMEM 1033 1321 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 455 462 ATP (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1321 AA; 146257 MW; 5443F48F79BF166 CRC64;

Query Match: 22.8%; Score 1472.5; DB 1; Length 1321;
Best Local Similarity 30.9%; Pred. No. 1.5e-87;
Matches 419; Conservative 234; Mismatches 530; Indels 171; Gaps 29;

QY 13 DGESESHPP--PKETEGSIS--DYLRIRYADKYDTLWIALICAGSGLPMSII 69
DB 23 DGSNNKRSRLDKMEGDIRVGFPELRFSSSKOILWLMGVCALLHMAQGGIIII 82
QY 70 FGSFKNYNNYNSGDSPE--AKAADOHFLWF----- 101
DB 83 FGIWTDIFIKDIQEROLEIPGKACVNTIVINSFHNWNGTCVGLVDISEMIKES 142
QY 102 --VLELGR--FVLYVSTAAITISAIRTRRLRVFLECTLRQDVVHFKQNGAIATYV 158
DB 143 GIYAGVTVILIGYFOIRLWITGARQRMRKIYFRIMRMEIGWFOCTSVGELNSRF 202
QY 159 TTNGRIOTGIAELVTVQALSNFSPAPVALASQWKLALIPMSYPAIFLVTCICAI 218
DB 203 ADDIEKINDATADOLAIQRMSTAMCGLLIGFYRGWKLTLVLAVSPLI----GIGA 258
QY 219 DAAQBARTRI-----YSRAAVLAEEVLSIRTVHAFYAQKWKYKDYDFLOQAHO 274
DB 259 IGLSTAKTELELKAAYAKAGSIADSVLSIRTVAAFGNGENKEVEREKNLVFAQRM 318
QY 275 SPNGVLFSTFYFCYAAIALAFKWFGRMYONGEVADVGKVTVLVSLVTLAATISMLAP 334
DB 319 GMYMGFFTYGMMCLIFCYALAPFWGSTLVLDSEETPGTLVQIFLCVTLAAMNIGH 378
QY 335 SGCVVYORRIFGSELSIDKPTQLDPLDPGKQPEGLQIBIQNLAFAPSPRSQA 394
DB 379 CLEIFSTGCSAATNIFOTIDRQVTDCKSGDGYKLDRIKGEIEFHNVTYHSPRDV 438
QY 395 RDNFLTIPAGKTTALVAGSGSKRMVGLLERWYLPSSGRILLDGLGLOYVWKLRLRI 454
DB 439 DNLKSVIAPGTTALVAGSSGAGKSTALQIORFYDPCEGMVTLGDHDIRSLNIRL 498
QY 455 RLVOQEPVLRFGTTFONANGFDEORLPREKQMLVOKACKASNGDVFINELPNGY 514
DB 499 GIVEQEPVLFSTTIAENIRFGED-----ATMEDIVQAADANAYNFINALPOQ 550

```

```

QY 515 EVGERAGALSGGORIATARSIIISOPKILLDEATSDALDPKAEKVQOEALNRVSKDR 574
DB 551 LVGEGGQMGSGQKQVAIARALIRNPKILLDMATSDALDNESEARVOEALNKIQHGTI 610
QY 575 LVTAHLKATVKSAGNIAVISOGKIVEGTHHELIIEFGCHYAALVRAQDLGADEOHEKT 634
DB 611 ISVAHRLSVRAADVIIGFEGHVAVERGTHEELLERKGVFMLVTLQSOG--ONAHK 668
QY 635 LHEKAAREAAAGERPALERTHTTA-----TSOAGDLKRRKVPVGTCTGYSLLKCI 690
DB 669 IMCKDATEGG-----FLERTPSGRYSRLRASIRQSKSOLSLTHOPPLAVA--- 720
QY 691 NLV-----WCFLA--STITVLIACAATFFCOALLFESRL 721
DB 721 SSYKDSKDNLDLVEEPAPVRRIKYNIPENWHYILVGLSAAINGAVTPYISLFSOLL 780
QY 722 TVFELSGHAAOERADFYILMEFFVVALG--NLGYFTTIGTCNVISOVVTHRYONAMP 779
DB 781 GTFSL--DKEOGRSEIHSKCLFFVILGCYSIPTQFLOGYTFKSCCELLTKRLKFGK 839
QY 780 LQODI-----ELLDIP-----EISGALTSQLSALPTQLOELISANFLIYVQ 824
DB 840 LGODIGNFDLNNPCVLTTLATDASQVQATGSOVGMVNVNFTNIIAALLIAFFSWK 899
QY 825 HRLEQCSSTLSWMEFGGCVWCTSTPAPCWLPPONSSRDEARSRLKGLCKRCKWACK 884
DB 900 LSL---IITIFFPLALSGAVTKMLTGF-----ASDQKALEKAGQITSEALSNI 948
QY 885 YRDPDLPLFDSRRPCSPYLVGHVEOGLAKIIQSF-----WFGRCFGFHL 928
DB 949 -----IVAGIGVEG--RPIKAFVELOTSYKTVARKNANIYGLCAF-- 987
QY 929 SQSMEFLAIAL-----GFCIADNVHLOVSTQTLNFISSSWAFCLPVOAAQYLA 983
DB 988 SOGIAFLAALAAKYGGFLIAIEG---LGFSHVFRVSVSA--LSATAVGTSTYTPSTA 1042
QY 984 KARSANYILMLRPTLKPIRETEKKGKPGVCGPVDLEDIEPRYRORDSARVLRGVSW 1043
DB 1043 KAKISARFPQLDRKPPINVSSEAGEKWNFGKIDIFDCKFTYPSRPIQVNLGLSV 1102
QY 1044 IEPQGVAVVAGSCGGKSTLIALSERPYDPTSGRISFAHENIAEMSPRLYGHMSLV 1103
DB 1103 VNPQTLAFVSSGCGKSTIQLLRYPDQGTVMIDCHSKRVNIQFURSNIQVSGE 1162
QY 1104 PTLYOGSVRENTVTLALA--ELSEELCOGRPARPMLWILSSLYOKALKRLAARQMG 1162
DB 1163 PVLFDCSINDIKYGDNTKISVERAIAAKQAQJHDFVMSLPKEYETNVGIO--G 1221
QY 1163 GORQIATARALIRNPKILLDEATSDALDPSERLVQAALDEASTSTTTTAVAHRL 1222
DB 1222 GEKQIATARAIVRDPKILLDEATSDALDSEKTVQALDKAREGRTCVIAHRLSTI 1281
QY 1223 NVDFIVFANGRIAGTTHAELQRLRGYYEMCL 1256
DB 1282 NSDIIAVVSGVVIETKTHEKLMARQAYYKLV 1315

RESULT 15
AB11_MOUSE
ID AB11_MOUSE STANDARD; PRT; 1321 AA.
AC 09QY30; Q9Q2E8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)
DE (Sister of P-glycoprotein).
GN ABCB11 OR BSEP OR SPGP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_Taxid=10090;
RP [1]
SEQUENCE FROM N.A.

```

RC TISSUE=Liver;  
 RX MEDLINE=20076398; PubMed=10607905;  
 RA Green R.M., Hoda F., Ward K.L.;  
 RT "Molecular cloning and characterization of the murine bile salt export  
 pump.";  
 RL Gene 241:117-123(2000).  
 RN [2]  
 RC SEQUENCE OF 463-635 FROM N.A.  
 RP STRAIN=C57BL/6; TISSUE=Liver;  
 RA Salker R., Suchy F.J., Ananthanarayanan M.;  
 RT "Molecular cloning of mouse liver bile salt export pump (bsep).";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
 CC INTO THE CANALICULUS OF HEPATOCYTES.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
 CC SITU.  
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES, AND  
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
 CC ATP BINDING CASSETTE (ABC) DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF133903; AAF14372.1; -;  
 CC EMBL; AF186585; AAD56419.1; -;  
 CC MGD; MGI:1351619; Abcb11.  
 CC InterPro; IPR003593; AAA.  
 CC InterPro; IPR001140; ABC\_transporter\_tmem.  
 CC InterPro; IPR003439; ABC\_transporter.  
 CC InterPro; IPR001687; ATP\_GTP\_A.  
 CC Pfam; PF00664; ABC\_membrane\_2.  
 CC Pfam; PF00005; ABC\_tran; 2.  
 CC SMART; SM00382; AAA; 2.  
 CC PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transmembrane; Transport.  
 FT DOMAIN 1 62  
 FT TRANSMEM 63 83  
 FT DOMAIN 84 147  
 FT TRANSMEM 148 168  
 FT DOMAIN 169 215  
 FT TRANSMEM 216 236  
 FT DOMAIN 237 240  
 FT TRANSMEM 241 261  
 FT DOMAIN 262 319  
 FT TRANSMEM 320 340  
 FT DOMAIN 341 353  
 FT TRANSMEM 354 374  
 FT DOMAIN 375 755  
 FT TRANSMEM 756 776  
 FT DOMAIN 777 794  
 FT TRANSMEM 795 815  
 FT DOMAIN 816 869  
 FT TRANSMEM 870 890  
 FT TRANSMEM 891 911  
 FT DOMAIN 912 979  
 FT TRANSMEM 980 1000  
 FT DOMAIN 1001 1011  
 FT TRANSMEM 1012 1032  
 FT DOMAIN 1033 1321  
 FT NP\_BIND 455 462  
 FT NP\_BIND 1113 1120  
 FT CARBOHYD 109 109  
 FT CARBOHYD 116 116  
 FT CARBOHYD 122 122

FT CARBOHYD 125 125 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 481 481 L -> P (IN REF. 2).  
 FT CONFLICT 633 633 T -> V (IN REF. 2).  
 SQ SEQUENCE 1321 AA; 146675 MW; 15B5EBF15D32967 CRC64;  
 Query Match 22.7%; Score 1468.5; DB 1; Length 1321;  
 Best Local Similarity 30.7%; Pred. No. 2.7e-87;  
 Matches 417; Conservative 232; Mismatches 530; Indels 179; Gaps 30;  
 QY 13 DG---SESSHPPEKTEGSIISDYLRIPRYADKYDWTNLNIALICAGSGASPLMSII 69  
 DB 23 DGHNDKXSRLODKKGEARVGFPELFRSSKDNWCMFPGSVGVCALLURHQAQGMIIIV 82  
 QY 70 FGSFTKNFNY-----NSGGSPSPAKADVHFVLWF 101  
 DB 83 FGLTDFEYDIEHQELSPGVCMNNTIWINSSFNONTNGT-SCGLVDINSEVIRK 141  
 QY 102 --VYLFIG--KEVLVYSTAAITSAITRTTLRRVFLLECTLRQEVNHFEDKOSGAIATX 157  
 DB 142 SGYAGVGVAVLILGFIQIRLWITGARIRKMKFYFRIRMRMEIGWDFCTSVGELNSR 201  
 QY 158 VTTNGNRIQTGIAELVFTVOALSMFSAFVVALASQMKLALITMSVIPAFLVFGICIA 217  
 DB 202 FSDINKIDEADOMALFLQRLSTALSGLLGFYRGMKLTLVILAVSPLI----GIGAA 257  
 QY 218 IDAAQEARITRI---YSRAVLAEEVLSIRTVHAFYAKKVKYKDYVFLQQAHQEKK 273  
 DB 258 VIGLSVAKTELEKAKAGSTADEVLSIRITVAAGGENKEVEREYKMLMFAQRGIW 317  
 QY 274 KSPNNGVLESTFYFCIYAALAFKFWGFRMYQNGEVADVGKVFVTVLSVLAATSISMLA 333  
 DB 318 KGMVGFFTGYMCLIFFCALAFWYGSRLVLDGEYTPOTLIQIFLCVILAAHMGAS 377  
 QY 334 PSGSVYQRRIFGSELSIDKPTOLDPLDPGSGKPECGIQIEIOLAFAYPSRPAQV 393  
 DB 378 SCLFESTGCSAASSIFQITIDROPVMDCMGSDGYKLDRIKGEIEFINVTFHPSPPEVKI 437  
 QY 394 LRDFNLTIPAGKTALVAGSGSKSTMVGLLRWYLPSSGRILLDGLGELGOYNVKKLSR 453  
 DB 438 LNNLSMVIKPGETTAFVSGSGAKSTALQIQRFYDPCGEMVTLDGHDIRSLNIEMLRDQ 497  
 QY 454 IRVQQEPVLFRTGTFQNIANGFMDEQRDLPREKQELVOKACKASNGOVFINELPNGYE 513  
 DB 498 IGIVEQEPVLFSTTTAENIRLG-----RESEATMEDIVQAARDANAYNFIMALPOQFD 549  
 QY 514 TEVGERAGALSGGQRIATARSIIISDPKILLIDRATSALDPKAEKVQAEALNRVSKORT 573  
 DB 550 TLVGEQGGOMSGGQKQVATARALIRKPKILLDLMATSDALDNESEAKVOCALNKIQHGT 609  
 QY 574 TLVIAHLATVKSAGNIATVISOQVIEQGTTHHELIEFGCHYAALVRAQOLGADQEQHEK 633  
 DB 610 IISVAHRLSTVRSADVIIIGFPHGTAVERTGHELLERKGVYFMLVTIQ---SQEDNTHKE 666  
 QY 634 TLHEKAAREAAAGERPALERTHTTATSQ---AGDLEKRVKVPVGTGLYSLIKGILLMPYEQ 689  
 DB 667 T-GIKGKDTTEGDTF--ERTFSRGSYQDSLRSIRQSKSQLSHLSHEPPLAI-----GDH 719  
 QY 690 KNLV-----WCFLL-STITVLICAAFPQALLFSRL 720  
 DB 720 KSSVDRKNDVLVEVEPAPVRRLIKYINISEWPYILVGCALCAAINGAVTPIYSLFSPQI 779  
 QY 721 LTVFELSGHAAOERADFYILMFFVVALG--NLVGYFTIGTWCNVISQVTHRYQAAMFOR 778  
 DB 780 LKTFSLV-DKEQORSEIYSMCLPFVILGCVSLTQFLQYFNFAKSGELTKRURKFGKA 838  
 QY 779 VLDDQTELLDPEQISGALTSOLSAPTOLOELISAN-----FLIVIVVQHR 827  
 DB 839 MLRQDIGNFDLKNPNPGLVTLTRLATDASQVQAGTSQVGMVNSFTNIEFVAVLAFLEPNW 898  
 QY 828 EOCSTSLAME-TGPGGCVMTSTPAFGMLPQNSRDEARSKLGLCKRCKWACKRSYR 886  
 DB 899 KLSIVISVFPFPLALSGAVQFKMLTGF-----ASODKEILEKAGQITNEALSNIIR---- 948

QY 887 DPORLIFDSRRPCSTVLGHVEOGLAKIIOSF-----WFGRCFGFHLSQ 930  
Db 949 -----TVAGIGVEG--RFKAFEVELEKSYKTAIRKANVYGLCYAF--SQ 989  
QY 931 SMEPLAIAL-----CFGIADNNLQVSTTOLNFISSWAF-----CLPVORAAQYLAYS 979  
Db 990 GISFLANSAARYGYLIVIED-----LAF---SYFVRVSSIAMSTAVGRTFSYT 1038  
QY 980 TSETKARSANYILRLPTIRETEENKKGPVCCPDLEIDIEPRYQORDSARVLRG 1039  
Db 1039 PSYAKAKISARFPOLLDRKPPIDVYSGAGERWDFQKIDFIDCKRTYPSRFDIQVLNG 1098  
QY 1040 VSWTIEPQPVAYVGCSCGKSTLIALSERFYDPTSGRISFAHENIAEMSPRLYRGHMSL 1099  
Db 1099 LSVSDPGQTLAFVSGCGKSTSIQLLERFYDQGTWMDGHDSKRVNVQPLRSNIGI 1158  
QY 1100 VQDEPTLYQGSVRENWTLALEA-ELSEELCOGRLPARPMLWLSSLYOKALKRLAAQROM 1158  
Db 1159 VSQEPVLFDCSINDNICYGDNTRISVERAIAAAQOLHDFVMSLPKRYETWVGIO-GS 1217  
QY 1159 QFSGCORIATARALIRNPKLLLDLDEATSDTOSERLYOALDEASTSRTTIAVAHRL 1218  
Db 1218 QLSRGEKRIATARAIVRDPKILLDEATSDTSEKTVQALDKAREGRTCIVIAHRL 1277  
QY 1219 STIRNVDFVFPANGRIAEOTHAELQRLAGRYEMCL 1256  
Db 1278 STIONSDIIAVWSQGVVIERGTHAKRLMDQNGAYIKLVI 1315

Search completed: April 27, 2002, 08:00:03  
Job time: 769 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 07:48:23 ; Search time 68.15 Seconds  
(without alignments)  
1780.790 Million cell updates/sec

Title: US-09-882-694A-11  
Perfect score: 6455  
Sequence: 1 MADESEKPRNQDGSSESSH.....LQRLGRYYECLQAQLDQA 1263

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*\n1: pir1:\*\n2: pir2:\*\n3: pir3:\*\n4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1970.5	30.5	1362	T41534	leptomycin B resis
2	1789.5	27.7	1408	T43261	multidrug resistan
3	1780	27.6	1307	T30882	multidrug resistan
4	1649.5	25.6	1280	DVHU1	multidrug resistan
5	1548.5	25.5	1287	S55692	multidrug resistan
6	1540.5	25.4	1279	DVHU3	multidrug resistan
7	1530.5	25.3	1276	DVMS1	multidrug resistan
8	1627.5	25.2	1276	A34786	multidrug resistan
9	1611	25.0	1276	DVHY1C	multidrug resistan
10	1601.5	24.8	1276	DVMS2	multidrug resistan
11	1601.5	24.8	1294	T19982	hypothetical prote
12	1594	24.7	1292	T48007	hypothetical prote
13	1592.5	24.7	1278	S41646	p-glycoprotein hom
14	1590.5	24.6	1281	T48123	p-glycoprotein iso
15	1590	24.6	1277	JH0502	p-glycoprotein - r
16	1585.5	24.6	1278	D86155	probable ABC trans
17	1556.5	24.1	1229	D85023	p-glycoprotein-lik
18	1556.5	24.1	1229	T52319	p-glycoprotein-lik
19	1554	24.1	1321	T23476	hypothetical prote
20	1553.5	24.1	1286	T02187	probable ABC trans
21	1551	24.0	1321	T27337	multidrug resistan
22	1549.5	24.0	1327	T21268	hypothetical prote
23	1544	23.9	1233	T04231	p-glycoprotein 2 -
24	1528.5	23.7	1266	T22090	hypothetical prote
25	1525.5	23.6	1230	D85023	probable P-glycopr
26	1525	23.6	1245	G86404	probable P-glycopr
27	1523.5	23.6	1283	A47377	multidrug resistan
28	1518.5	23.5	1229	F86155	probable ABC trans
29	1517	23.5	1275	T31073	multidrug resistan

30	1512	23.4	1302	2	S30327	multidrug resistan
31	1507.5	23.4	1302	2	A41249	multidrug resistan
32	1502	23.3	1268	2	T22094	hypothetical prote
33	1495	23.2	1286	2	A42150	p-glycoprotein ppp
34	1494	23.1	1270	2	T21269	hypothetical prote
35	1488.5	23.1	1289	2	D87789	protein C34G6.4 [1
36	1473.5	22.8	1104	1	DVMS1A	multidrug resistan
37	1472.5	22.8	1321	2	T28842	bile salt transpor
38	1471.5	22.8	1322	2	T06165	multidrug resistan
39	1470.5	22.8	1321	2	T42228	probable P-glycopr
40	1470	22.8	1247	2	F86405	hypothetical prote
41	1468.5	22.7	1291	2	T21267	hypothetical prote
42	1467.5	22.7	1323	2	H85202	hypothetical prote
43	1464	22.7	1318	2	T21266	hypothetical prote
44	1457	22.6	1310	2	S30328	multidrug resistan
45	1454.5	22.5	1254	2	T30855	multidrug resistan

ALIGNMENTS

RESULT 1

T41534 leptomycin B resistance protein, ABC transporter [imported] - fission yeast (Schizos

N:Alternate names: p-glycoprotein homolog

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Feb-2001

C:Accession: T41534; S20548

R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1998

A:Reference number: Z17318

A:Accession: T41534

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1362 <SEE>

A:Cross-references: EMBL:AL031307; NID:g3426126; PIDN:CAA20363.1; PID:g3426129; GSPD

A:Experimental source: strain 972h; cosmid c663

R:Nishi, K.; Yoshida, M.; Nishimura, M.; Nishikawa, M.; Nishiyama, M.; Horiouchi, S

Mol. Microbiol. 6, 761-769, 1992

A:Title: A leptomycin B resistance gene of Schizosaccharomyces pombe encodes a prote

A:Reference number: S20548; MUID:92244046

A:Accession: S20548

A:Molecule type: DNA

A:Residues: 1-1335, TC, 1336-1362 <NIS>

A:Cross-references: EMBL:U10695; NID:g218549; PIDN:BAA01537.1; PID:d1002012; PID:g21

C:Genetics:

A:Gene: SPDB:SPCC663.03; pmd1

A:Map position: 3

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homol

C:Keywords: ATP; glycoprotein; membrane protein; nucleotide binding; p-loop

F:438-641/Domain: ATP-binding cassette homology <ABC1>

F:455-462/Region: nucleotide-binding motif A (p-loop)

F:1137-1332/Domain: ATP-binding cassette homology <ABC2>

F:1154-1161/Region: nucleotide-binding motif A (p-loop)

Query Match 30.5%; Score 1970.5; DB 2; Length 1362;

Best Local Similarity 35.5%; Pred. NO. 4.2e-133;

Matches 484; Conservative 225; Mismatches 508; Indels 147; Gaps 24;

Qy 9 RPNDGSES-----SSHPPEKE-----TEGSISDYLRIFRYADKVDWFLNV 50

Db 37 KSENEASESHVVDVKKDPPEQVTPPEOELIKQINDTPAKLSGYPRILSYADKWDIMQL 96

Qy 51 IALICATGSCASPLMSIIFGSGTNKFNYSNGDGSPEAFKADVDHFWLFWFLVFLGKPV 110

Db 97 AGTITGAGGMPMSLVSGQLAQAFTDLASGKA-SFOHTVDHFLCLFIYIAIVGFC 155

Qy 111 LTVYSTAAITISARTTRTLRRVFLCTLRQVHFDKQSNIAIATXVTYTNNGRIQFGIA 170

Db 156 CSVIYVTFYTIAGRIARRIRQDYLHAILSONIGYFDBLGAGEITTRITTTDTPNF 215

Qy 171 EKLVTFTQVLSMFFSFAVVALASQWKIALITMSVIPAFILVLTGICIAIDAAQEARITRY 230





```

Qy 887 DPORLIFDS-----RRPCSPVIVLGHVQGLAKI1QSPWRCRCFGPHLSQSMFLA 936
Db 932 TVVSLMETEALQSQAQLRQLKSDILPIVKSLL-----YASSQALPFFC 978
Qy 937 IALGFCIAVDNML-----QVSTTQLNFISSWAFCLPVQAAOYLAVSTSFTRKASA 988
Db 979 MALGP-----WYGGSLHGGEYSLFDFYVCFSEVIF--GAQAAGTFFSHAPDMGKAKIA 1030
Qy 989 ANYILMLRLKPTIRE--TEENKKG-PVGGCP--VLEDEFFRYRQRDSARVLGVSMT 1043
Db 1031 A-----REFKRLFSSTGMASRSGKVPVTSMRGLVEFRDVSFRYPSRLQEPILRLNLT 1084
Qy 1044 IEPQGVAVVAGSCGSLTIALISFYDPTSGRISFAHENAIDMSPLRYGHMSLVQOE 1103
Db 1085 IKPGQFVALVAGSGSGSTIALIERFYDPLKGGVYVQDKNIITLSSSYRSHALISOE 1144
Qy 1104 PLYGSGVRENVTLALEAE--LSBELCOGRLPRLPMLWLSLYQKALKRLAAQRMQFIS 1161
Db 1145 PTLFOGTIRENTILLGSPNPHWTDVFLVKACKDANIYDFILS--LPQGFNTIVGNKGMLS 1202
Qy 1162 GGQRQIARALARNPKLLLDLDEATSLDQSERLVOAALDEASTSRTTIAVAHRLSTI 1221
Db 1203 GGQKQIARALARNPKILLDEATSLDSESENVVQAAALDAARGRITIAVAHRLSTI 1262
Qy 1222 RNVDFIVFANGRIAETGTHAEQLRLGRYEMCLQAQLD 1261
Db 1263 QRADLIYVLDGVEVSGTHRELLRKKGRYELVHLQNPD 1302

RESULT 4
MDVHUI
multidrug resistance protein 1 - human
C:Alternate names: P-glycoprotein 1
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990; sequence revision 18-Aug-1995; #text_change 19-Jan-2001
C:Accession: A34914; PS0162; S15500; A25059; S43838; I52236; I65204
R:Chen, C.; Clark, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
J. Biol. Chem. 265, 506-514, 1990
A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
A:Reference number: A34914; MUID:90094448
A:Accession: A34914
A:Molecule type: DNA
A:Residues: 1-1280 <CHE>
A:Cross-references: GB:M29447; GB:J05168; NID:g187496; PIDN:AAA59576.1; PID:g386862
Submitted to JPIB, April 1991
A:Reference number: PS0162
A:Accession: PS0162
A:Molecule type: DNA
A:Residues: 1-22 <KIO>
R:Kioka, N.; Yamano, Y.; Komano, T.; Ueda, K.
Submitted to the EMBL Data Library, April 1991
A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) expression
A:Reference number: S15500
A:Accession: S15500
A:Molecule type: DNA
A:Residues: 1-22, 'R' <K12>
A:Cross-references: EMBL:X58723; NID:g34522; PIDN:CAA41558.1; PID:g34523
R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I.
Cell 47, 381-389, 1986
A:Title: Internal duplication and homology with bacterial transport proteins in the mdrl
A:Reference number: A25059; MUID:87028230
A:Accession: A25059
A:Molecule type: mRNA
A:Residues: 1-184, 'V', 186-1280 <CH2>
A:Cross-references: GB:M14758; NID:g187468; PIDN:AAA59575.1; PID:g307180
R:Chambers, T.C.; Pohl, J.; Glass, D.B.; Kuo, J.F.
Biochem. J. 299, 309-315, 1994
A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase of
A:Reference number: S43838; MUID:94220047
A:Accession: S43838
A:Molecule type: protein
A:Residues: 656-689 <CHA>

```

R:Gekeler, V.; Weger, S.; Probat, H.  
 Biochem. Biophys. Res. Commun. 169, 796-802, 1990  
 A:Title: mdrl/p-glycoprotein gene segments analyzed from various human leukemic cel  
 A:Reference number: I52238; MUID:90290529  
 A:Accession: I52238  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 178-215 <RES>  
 A:Cross-references: GB:M37724; NID:g183537; PIDN:AAA8047.1; PID:g553314  
 A:Accession: I65204  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 800-856 <RE2>  
 A:Cross-references: GB:M37725; NID:g183538; PIDN:AAA8048.1; PID:g553315  
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant  
 tractually and functionally unrelated lipophilic anticancer drugs.  
 C:Genetics:  
 A:Gene: GDB:PGY1; MDR1  
 A:Cross-references: GDB:120712; OMIM:171050  
 A:Map position: 7q21-7q21  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotei  
 F:1-638,653-1280/Region: duplication  
 F:1-638,653-1280/Region: duplication  
 F:49-350/Domain: hydrophobic <HBL>  
 F:351-637/Domain: hydrophobic <HBL>  
 F:410-604/Domain: ATP-binding cassette homology <ABC1>  
 F:427-434/Region: nucleotide-binding motif A (P-loop)  
 F:551-555/Region: nucleotide-binding motif B  
 F:638-708/Domain: linker <LIN>  
 F:709-993/Domain: hydrophobic <HB2>  
 F:984-1280/Domain: hydrophobic <HB2>  
 F:1053-1249/Domain: ATP-binding cassette homology <ABC2>  
 F:1070-1077/Region: nucleotide-binding motif A (P-loop)  
 F:1196-1200/Region: nucleotide-binding motif B  
 F:91,94,99/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:433/Binding site: ATP (Lys) #status predicted  
 F:661,667,671/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status  
 F:667,671,683/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #st  
 F:1076/Binding site: ATP (Lys) #status predicted

Query Match 25.6%; Score 1649.5; DB 1; Length 1280;  
 Best Local Similarity 32.3%; Pred. No. 5e-110;  
 Matches 428; Conservative 243; Mismatches 490; Indels 163; Gaps 26;  
 Qy 24 EKETEGSISDYLRIPRYADKYDWTNLVNLICAIAGSGASLPMSIIFSGFTNFNN--- 79  
 Db 27 KKEKFTVSVP-SMFRYSNMLDKLYMVVGTAAIHHGAGLPMLMLVFGEMTDIFANAGNL 85  
 Qy 80 -----YNSGDGSPFAF---KADVDHFVLMFVFLFGKFLVTVVSTAATITSAINTTR 128  
 Db 86 EDLMSNITHRSNDINDTGFMMLEEDMTRYAYVYSGIGAGLVAAVYQVSWFCLAAQRQIH 145  
 Qy 129 TLRRVFECLTRQEVNHFQKQNGAIATATVYTNNGNRIQTGAELVFTVQALSMFFSAPV 188  
 Db 146 KIRKQFFHAIMROEIGCNFVDHVGELNTRLTDDVSKINEGIDGKICMFPFSMATFTFTGI 205  
 Qy 189 VALASQWKLALITMSVIPAIFLVTVGICIAIDAAQAEARTIRYSRAAVLAAEVLSSIRTVH 248  
 Db 206 VGFTRGWKLTLLVILASIVPLGSLAAVYNAKILSSFTDKELLAYAKACAVAAEVLAAITVI 265  
 Qy 249 AFYAQKKMYEKYDVPVLPQAHQEGKKKSPNNGVLFSTPEFYCIAAATALAFWKGRFMYQNGE 308  
 Db 266 AFGGQKKELERYKNKNLEAKRIGIKKAITANISGAALFLIYASVALAFWYGTTLVLSGE 325  
 Qy 309 VADVGKVFVTVLSVTLAATISIMSLRSGSVVYORRIFGSELSFIIDKPTOLDPLDPSGQK 368  
 Db 326 YS-IGQVLTVFVSVLIGAFVSGQASPSIEAFANARGAAVEIFKIIDNKPISIDYSKSGHK 384  
 Qy 369 PEGCLGOIEIQLNLAIPAYRPSAQVRLDFMLTIPAGKTTALVAGSGKSGKSVGLLERYW 428  
 Db 385 PDNIKGNLEFRNVHPSYPSRKYVKILGNLKVQSGQTVLVAGNCGCKSTTVQLMQRLY 444



```
QY 429 LPSSGRILLDGLGOYNVYKWLRSRLRVQPEPVLFRGTIFONIANFGFMDEQDRLPREK- 487
Db 445 DTEGMSVDSQDQIRTNIRFLREILGVVDSQEPVLFATTIAENIRY-----RENV 495
QY 488 OMELVOKACASGVDFINELNGYETVEVRGAGALSGGQORORAIARISIIISDRKILLD 547
Db 496 TMDEIEKAVENAYDFIMKLKFDFTLVGERGALSGGQORORAIARALVKNPKILLD 555
QY 548 EATSALDPKAEKVQALNRVSKDRTTLVIAHKLATVKSAGNIATVISOQKIVEQTHHEL 607
Db 556 EATSALDTEAEVAVQALDKARKGRTTIVIAHRLSTVRNADVTAGDGVIVKRGHDEL 615
QY 608 IFECHYAALVAQDLGADQEQHEKTLHEKAAREAAAGERPALERT----- 653
Db 616 MKERGIYFKLVNTQAGNEVELEN-----AADESKSEIDALEMSNDSRLIRKRST 668
QY 654 -HTTATSQAGD-----LEKKKVPVGTGLSYLLKCILIMFYEORNLVW-CFLLTITV 703
Db 669 RRSVRGSAQDRKLSLTKALDESIPVSEF--WRIMKLNL-----TEMPYFVGVFCA 718
QY 704 LICAATFPGOALLFSRLTVFEL--SGHAAOERADVFILMFFVVALGNLVGYFTTGWCN 761
Db 719 IINGLOPAFAIIPSKIIGVFRIDDPETKRONSLFSLFLALGIIIPITFLQGFPG 778
QY 762 VISQVTHRYQAAAMFORVLDODIELDIPQISALTSOLSALPTQLOELISANFLIIV 821
Db 779 KAGELTKURYVYFMSLRQDQVSWFDDPKNTTGALTTRLANDAAQVKAIGSLAV--- 835
QY 822 VQGHLEQCSTTSLSMETPGGCVNCTSTPAFGW-----LPQ 858
Db 836 -----ITQNIANLGTGII-----SFTYGWQLTLALLATVPIIAIAGVYENKMLSG 881
QY 859 NSSRDEARSKLGLCKRCKWACKRSYRDPDLIFDSRRPSCPVTVLGHVQGLAKIIQ-- 916
Db 882 QALKDKKELGAGKIAIEA-----IENERTVYSLTQDKPEHYAQSLOVP 927
QY 917 -----SFWFGRCFGHLSQMEFLAIA-----LGFCIAVDNNMLVSTTQLNFISSNAF 965
Db 928 YRNSLRKAHLEFGITFSF--TQAMMYFSYAGCFREGAYLVAKHKLMSFEDVLLVS---AV 981
QY 966 CLPVOAAQYLAYSTFTKARSANVILWRLTKPTI--RETEENKKKPGVCGCPVDLE- 1022
Db 982 VFGAMAVQVSSFAPOYAKAKISAHHIIMIEKTPLDISYSTE-----GLMENTLEG 1033
QY 1023 -----DIEFRYRORDSARVLGVSMTEPQGQVAYVAGSCGKSTLIALSERGYDPTSGR 1077
Db 1034 NVTCEVFNYPTRPDIPVQGLSLEVRAGQTLALVSSGCGKSTVQVLLERFYDPLAGK 1093
QY 1078 ISPAHENIAMSRLYRGMSLVQOEPPLYOGSVRENTVIALAEEL--SEELCQGRLPAR 1135
Db 1094 VLLDGEKEIKRLNVOMLRAHLGVSQEPILFDCSIAENIAYGDNRSRVSGEEIVRAAKAN 1153
QY 1136 PMLMILLSLYOKALKRLAARGMOMFGSGGQORORAIARALIRNPKLLLDLEATSDTQSE 1195
Db 1154 IHAFIESLPKNYSK--VGDKGTLQSGGQORORAIARALVROPHIILLDEATSDTQSE 1211
QY 1196 RLVOAALDEASTRTTIVIAHRLSTVRNADVIFVFAANGRTAETGTHAEQLRLRGYEMC 1255
Db 1212 KVVGEALDKAREKTCIVIAHRLSTIQNADLIIVVFQNGRVKHEGTHQOLLAQKGIYSMV 1271
QY 1256 LAQS 1259
Db 1272 SVQA 1275
```

## RESULT 5

S55692

multidrug resistance protein homolog (mdr) - African clawed frog

C:Species: xenopus laevis (African clawed frog)

C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 02-Feb-2001

C:Accession: S55692

R:Castillo, G.; Shen, H.J.; Horwitz, S.B.

Biochim. Biophys. Acta 1262, 113-123, 1995

A:Title: A homologue of the mammalian multidrug resistance gene (mdr) is functionally  
A:Reference number: S55692; MUID:95322451  
A:Accession: S55692  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1287 <CAS>  
A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:420-614/Domain: ATP-binding cassette homology <ABCI>  
F:437-444/Region: nucleotide-binding motif A (P-loop)  
F:1062-1258/Domain: ATP-binding cassette homology <ABC2>  
F:1079-1086/Region: nucleotide-binding motif A (P-loop)

Query Match 25.5%; Score 1648.5; DB 2; Length 1287;  
Best Local Similarity 32.3%; Pred. No. 5.9e-110;  
Matches 429; Conservative 231; Mismatches 523; Indels 145; Gaps 24;

```
QY 10 PNODGSE-----SSSHPPPEKETEGSISDYLRIFRYADKYDTLNVIALICAGISGLP 64
Db 22 PNSNREKKGPFSEKFKKKKETEPPKVGVTMFRKYSTSDSKMLMFGTILASLHGAALP 81
QY 65 LMSLIIGSTNKNFNNVNSD-----GSPFAKADVDHVLFWFVFLFGICGFVLIYV 114
Db 82 LMLVLFGEMTDFVNVNQVDYTGNTFWESMINASRELOQMITYAYIYISGLGFGVMCAIY 141
QY 115 STAAITISARTTRTLRRVFLUETLRQEVWHFDPKQSNCAIATXVTNNGRNIQTGIAEKLV 174
Db 142 QISFWTLRAGROIKIRSNFFHVRQEIWQFIDNAGELNRLTDDVSKINEGIGKIA 201
QY 175 FTQALSMFSAFVVALASOMKALITMSVIPAFILVTGICIAIDAAQEARITRIYSRAA 234
Db 202 MLLQSLTTLVTGFIIGIKGWLKLTWVMGATSPINGLSAAIWAARVLSAFTNKLKAYAKAG 261
QY 235 VLAEEVLSRTVTHAFYAGKMKVRYDFLQQAQEGKKSPNNGVLFSTVFYCIYAAIA 294
Db 262 AVAEEVLSRTVTHAFYAGKMKVRYDFLQQAQEGKKSPNNGVLFSTVFYCIYAAIA 321
QY 295 LAFWKCFRMYONCEVADGVKFTVSVTLAATSISMLAPSGSVYVYORRIFSGSELSIID 354
Db 322 LAFWYCTLLIDGGYT-IGSVLTVPFAVILGAFAGVGTSENIEAFANARGAAYTIFNID 380
QY 355 KPTQLDPLDPSKOPGECGLGOIEITONLAFAYRPSQVLRDQENLTIPAGKTTALVGASG 414
Db 381 NQPKISFSGKEGLPKDKINGDIEKKNVIFYPSKQIQVQLKGLNLTIPSGRTVALVSSG 440
QY 415 SGKSTWVGLLERWYLPSSGRILLDGLGELGQVNVKWLRSRLRVQPEPVLFRGTIFONIAN 474
Db 441 CGKSTTVQLQIRFYDPDEGVITLDGQDIRSINIRYDREIIGVVSQEPILFDTTADNIRY 500
QY 475 GPMDEQRDLPREKOMBELVOKACASNGSDVFTNELPNGYETVEVRGAGALSGGQORORAI 534
Db 501 G---REDVTKEE---TERATKEANAYDFTMKLPKLETLVGERGTQLSGGQORORAI 552
QY 535 RSTISDPKILLDEATSDALDPKAEKVQAEALNRVSKDRTTLVIAHKLATVKSAGNIATV 594
Db 553 RALVRNPKILLDEATSDALDTESEAVVQSALDKAREGRTTIVVAHRLSTIRNAAIGFD 612
QY 595 QGKIVSQGTHHELIEFGCHYAALVAQDLGADQEQE-----HEKTL-----HEKAARE 642
Db 613 NGVIVEGSGSKELMERGGVYENLVTLQVETSKDTEEDLETHIYEKKIPVTHTHSNLVR 672
QY 643 AAGERPALERTHTTATTSQAGDLKXKVPVGTGLSYLLKCILIMFYEORNLVWFLSTIT 702
Db 673 KSSRNTIKSVETEDKEVDEEEKK--EEGPPVPVFFKVNKL-----NKPEYFVVGVC 727
QY 703 VLICAAITFGQALLFSRLTLVTFELSGHAAOERAD--FYILMFVVALGNLVGYFTTGWC 760
Db 728 AMINGATQPAFIIFSRIGVFF--AGPVSQMRSESSYLLFLALGGVSTIFFFLQRTFE 785
QY 761 NVISQVTHRYQAAAMFORVLDQDIELDIPQISALTSOLSALPTQLOELISANFLIIV 813
Db 785 NVISQVTHRYQAAAMFORVLDQDIELDIPQISALTSOLSALPTQLOELISANFLIIV 813
```

Db 786 KGAEITLMLRLGSGFKMLRQIEGWFDSDSKNSTGALTRTLATDASOVQATGTRIALLA 845  
Qy 814 AN-----FLIVVYQHRLEQCTTSLMNETGPGCVWCTSTPAFGWLPQNSSRDEAR 866  
Db 846 ONVANLGTATIIISFIYG-----WQTLTLLIATVP--VIAAGLVEMKMPAGHAKDKKE 897  
Qy 867 SRKLGKLC-----RKWACKRSRYRDPDRILFDSRRPCCSPTVLGHVEQG 910  
Db 898 LEKAGKISTDAVINIRTVVSLTRERKFEAMYEKSLGEPYR----- 937  
Qy 911 LAKIIQSFWRGRCFGPHLSQSEFLAI-----ALGFCFIVDNMLQVSTTQNLFISSWAF 965  
Db 938 -NSIKKAHLHGLTYG--LSOAHVLCWVSVLGAVLVLEGLMKDEV---FLVSS-AI 990  
Qy 966 CLPVOAAQVLAYSTFTKARSANVILWRLTKPTRETEENKKKGPVGGCVDLEDE 1025  
Db 991 VLGAALGQTSFAPQYTKAMISAMHIFSLLEVPQIDSYSDQEKPKNCNGNVFKGVN 1050  
Qy 1026 FRYQRDSARLVKGSWKTIEPGQVAYVYAGCGCKSTLIALSERFPTDPTSGRISFAHEH 1085  
Db 1051 FNYPTREDITVLQGLDLSVQGETLALVSGCGCKSTVSLERFYDPFEGEVLVDGLSV 1110  
Qy 1086 AEMSPRYRGMHSLVQOEPYLYGGSVRENV-----TLALEAESEBELCG 1130  
Db 1111 RNLIQVRAQMGIVSOEPILEDCSIGDNIAYGDNRRKVTQEEETAAKEANTHS----- 1165  
Qy 1131 RLPAEPMILSLSSYLKALKRLAAQGMQFSGGQRQRIATARALIRNPKILLDEATSL 1190  
Db 1166 -----FIESLTDKYNTRV-GDKGTQLSGGQKQRIATARALIRNPKILLDEATSL 1215  
Qy 1191 DTQSERVQAALDEASTSTTIAVAHRLSTIRNVDPVFPVAFNGRIAGTGPHELOLRGR 1250  
Db 1216 DTESEKVVQALDKARMGRTCVIAHRLSTIQADKIAVQNGKVVYEQGTHQQLQKGV 1275  
Qy 1251 YXENCLAQ 1258  
Db 1276 YFSLVTIQ 1283

## RESULT 6

DVHU3  
Multidrug resistance protein 3 - human  
N:Alternate names: P-glycoprotein MDR3  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
C:Accession: JS0051; S01346; A42213; I38015  
R:van der Bliek, A.M.; Koeliman, P.M.; Schneider, C.; Borst, P.  
Gene 71, 401-411, 1988  
A:Title: Sequence of mdr3 cDNA encoding a human P-glycoprotein.  
A:Reference number: JS0051; MUID:89138016  
A:Accession: JS0051  
A:Molecule type: mRNA  
A:Residues: 1-1279 <VA1>  
A:Cross-references: GB:M21334; NID:g187501; PIDN:AAA36207.1; PID:g307181  
R:van der Bliek, A.M.; Baas, F.; ten Houte de Lange, T.; Koeliman, P.M.; van der Velde-ke  
EMBO J. 6, 3325-3331, 1987  
A:Title: The human mdr3 gene encodes a novel P-glycoprotein homologue and gives rise to  
A:Reference number: S01346; MUID:86111519  
A:Accession: S01346  
A:Molecule type: mRNA  
A:Residues: 856-1093, 'FVDFGFGQ', 1094-1279 <VA2>  
A:Cross-references: EMBL:X06181; NID:g34524; PIDN:CAA29547.1; PID:g34525  
R:Lincke, C.R.; Smit, J.J.M.; van der Velde-Koerts, T.; Borst, P.  
J. Biol. Chem. 266, 5303-5310, 1991  
A:Title: Structure of the human MDR3 gene and physical mapping of the human MDR locus.  
A:Reference number: A42213; MUID:91161629  
A:Accession: A42213  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-1279 <LIN>  
R:Smit, J.J.; Mol, C.A.; van Deemter, L.; Wagenaar, E.; Schinkel, A.H.; Borst, P.  
Biochim. Biophys. Acta 1261, 44-56, 1995  
A:Title: Characterization of the promoter region of the human MDR3 P-glycoprotein gene.

A:Reference number: I38015; MUID:95200972  
A:Accession: I38015  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-61, 'RGSSRDVLAQC' <RES>  
A:Cross-references: EMBL:235284; NID:g1006662; PIDN:CAA84542.1; PID:g1006663  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistar  
tructurally and functionally unrelated lipophilic antitumor drugs.  
C:Genetics:  
A:Gene: GDB:PGY3; MDR3  
A:Cross-references: GDB:I20713; OMIM:171060  
A:Map position: 7q21-7q21  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; Transmembra  
F:1-640/53-1279/Region: duplication  
F:1-606/Domain: ATP-binding cassette homology <ABC1>  
F:429-436/Region: nucleotide-binding motif A (P-loop)  
F:553-557/Region: nucleotide-binding motif B  
F:638-694/Domain: linker <LINK>  
F:1052-1148/Domain: ATP-binding cassette homology <ABC2>  
F:1069-1076/Region: nucleotide-binding motif A (P-loop)  
F:1193-1199/Region: nucleotide-binding motif B  
F:91,97/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:435/Binding site: ATP (Lys) #status predicted  
F:1075/Binding site: ATP (Lys) #status predicted

Query Match 25.4% Score 1640.5; DB 1; Length 1279;  
Best Local Similarity 33.2% Pred. No. 2.2e-109;  
Matches 422; Conservative 230; Mismatches 532; Indels 89; Gaps 21;

Qy 35 LRIFRYADKYDWTLVNIALICAGSGASLPLMSIIFGFTNKNFNPNVNSGDGSEAF---- 90  
Db 43 LTLFRYSQWQDKLPMSLGTINAIAGHSGSLPLMIVFGEMTDKFDVDTAGNFSPVFNLSL 102  
Qy 91 -----KADVDHFLVFLVFLVFLVFLVFLVFLVFLVFLVFLVFLVFLVFLVFLVFLV 143  
Db 103 LAPGKLEEMTRVAYYISGLGAGVLAIVQSVFTLAAGROIRKQKFFHALLRDEI 162  
Qy 144 WHFDKQSGATATXVTNGNRIOGTIAEKLVFTVQALSMFFSAFVVALASQKHALIWS 203  
Db 163 GWFDINDTTELMTRLTDIISKISGIDGKVMGFFQAVATFFAGFVGTFRGKRLVITMA 222  
Qy 204 VIPAFLVTVGICIAIDAAQAEARITRIYSRAVLAEEVLSSIRTVIAFYAQKMKVEYDF 263  
Db 223 ISPIGLSAAVWAKILSAFSDKEALAAAYAKAGAAEALGAIRTVIAFGQNKELERYOKH 282  
Qy 264 LQQAQHEGKKSPNNGVLFSTEYFCIYAAIALAFWKGFMYQNGEADVGVKVTYVLSVT 323  
Db 283 LENAKEIGIKKAISANISMGIAFLIYASVALAFWYGLVLSKEVT-IGNANTVFFSIL 341  
Qy 324 LAATSIKSLAPSGSVVYORRIFGSELSIIDKPTOLDPLDPSGKQPEGCGIETONLAF 383  
Db 342 IGAFSGQAAPCIDAFANARGAAVYFIDIDNNPKIDSEFSEHGPDSIKGNLEFNDVHF 401  
Qy 384 AVPSRPSAOVLDRPNTIPACKTTALVAGSGCKSTMVGLIERWYLPSSGRILLDGLGLG 443  
Db 402 SPSRANVILKGLNKLAVQSGCTVALVSGCGKSTTVQLIQRLYDPDETINIGQDIR 461  
Qy 444 QYNVWKLRSIRLVOQEPVLFRTIFONIANG-----FMDGORDLPREKOMEVLVQKACAS 499  
Db 462 NFNYNTRIEIGWVSQEPVLFSTTIAENICYGRGNVTMD------IKKAVREA 509  
Qy 500 NGDVFTELNPNGYETVEGERAGALSGGQRIARIAHSIISDPKILLLLDEATSDALDPKAEK 559  
Db 510 NAYEETMKLPQKFDPLVNGERAGLQSGQKQRIARIAHALVRNPKILLLLDEATSDALDESEA 569  
Qy 560 VYQEAALNRVSKDRTPVIAHKLATVKAGNIAVISQGIKEVQCTHHELFECFCHYAALVR 619  
Db 570 EVOAALDKAREGRTTIVIAHRLSVRNADVAGFEDCVIEQGSHELSMKKGVYFKLVN 629  
Qy 620 ADDLGADEQOEHEKTLHEKAAREAA--GEAPALERHTTATSOAGCLEKRVKVPVGTGL-- 675

Db 630 MOTSGSOISEEFELNDEKAAATMAPNGWKSRLFRHSTOKNLKNSOMCOKSLDVEDTGLE 689  
Qy 676 -----YSLKCLILMEYOKLYMFCFLSLITVILICAATPPGOALLFSRLTLVFELSGHA 730  
Db 690 ANVPVPVFLKVLK-----NKTEWPFVVGTVCAIANGGLQAPAFSVTFSEIIAIFPGGDDA 745  
Qy 731 A-QERADVFLMEFVVALGNLVGFTIGWTCNVISOVVTHRYOAAAFORVLDLLELDI 789  
Db 746 VKQCKNIFSLIFLGIISFTTFELOGFTFGKAGELLIRLRRSNMAKRLRQDSMFDD 805  
Qy 790 PEQISGALFOLSALPTOLO-----ELISAN-----FLIYVWGHRLEQCSITSL 835  
Db 806 HKNSTGALSRLIATDAQVQATGTRIALAQNLANLGTGIIISFYG-----WQLTLL 859  
Qy 836 WMETGGCGVWCVTSTPAGFWLPQNSRDDEARSKLGLCKRCKWACKRRSYRDPDKLIPDS 895  
Db 860 LLAVVP--IIAVSGVIVEMKLLAGNAKDKKELEAACKIATEAIENIRTVVSIQERKFES 917  
Qy 896 RRPCSPTVLGHVREQGLAKIIQSFWGRCFGFHLSQSNFEFLATALGFCIAVDNWLQVSTQ 955  
Db 918 M-----YVEKLYGPRVRSVQKAHIYGITFSISOAFMYFSYACGFRFGAYLIVN-GH 967  
Qy 956 LNFITSSWAFCLPVQAAAQYLAYSF-----TKARSAANYILWLRTLKPTIRE-TEENKK 1010  
Db 968 MRFRDVLIVSAIVGAVA-LGHASSFAPDYAKAKLSAHLFWLPERQPLINDSYSEGLK 1026  
Qy 1011 KGPVGGCPDLEIDIEFRYQRDSARVLGVSWTIEPGQFVAYVYGASGCKSKTLIALSERF 1070  
Db 1027 PDKEG-NITFEVFNPTFRANFVLQGLSLEVKKGOTIALVGGSGCKSTVVGOLLERF 1085  
Qy 1071 YDPTSGRISPAENIAEMSPRLYRGHMSLVQOEPTLYQGSVRENVTPLAEAL--SEELC 1128  
Db 1086 YDLPLAGTVLLDGOEAKLNVQWKLRAQLGIYSQEBPILFDCSIAENIAYGNSRVWSQDBIV 1145  
Qy 1129 QGRIPA--RPMILWSSLYOKALKRLAAQRMQFSGGQORQIAIRAIIRNPKLILLDEA 1186  
Db 1146 SAAKAANIHPFIETLPHKYETRV---CDKGTQLSGGQKQRTAIARALIRQPQIILLDEA 1201  
Qy 1187 TSALDTQSRRLVQALDEASTRTTIAVAHRLSTIRNVDFVIFVANGRAETGTHAEIQR 1246  
Db 1202 TSALDTSEKVVQZALDRAREGTCIVTAHRLSTIQNALDILVVFQNGRYKEHGTHQQLA 1261  
Qy 1247 LRGRYEMCLAQS 1259  
Db 1262 QKGIYFSWVSVQA 1274

RESULT 7  
DVM51  
multidrug resistance protein 1 - mouse  
N:Alternate names: p-glycoprotein 1  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
A:Accession: A33719; A25057; I57510  
R:Raymond, M.; Gros, P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 6488-6492, 1989  
A>Title: Mammalian multidrug resistance gene: correlation of exon organization with stru  
A:Reference number: A33719; MUID:89367274  
A:Accession: A33719  
A:Molecule type: DNA  
A:Residues: 1-1276 <RAY>  
R:Gros, P.; Croop, J.; Housman, D.  
Cell 47, 371-380, 1986  
A>Title: Mammalian multidrug resistance gene: complete cDNA sequence indicates strong ho  
A:Reference number: A25057; MUID:87028229  
A:Accession: A25057  
A:Molecule type: mRNA  
A:Residues: 1-1276 <GRO>  
A:Cross-references: GB:M14757; NID:g199100; PIDN:AAA79005.1; PID:g387426  
R:Raymond, M.; Gros, P.  
Mol. Cell. Biol. 10, 6036-6040, 1990  
A>Title: Cell-specific activity of cis-acting regulatory elements in the promoter of the  
A:Reference number: I57510; MUID:91042535

A:Accession: I57510  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-21 <RES>  
A:Cross-references: GB:M60348; NID:g199102; PIDN:AAA39513.1; PID:g554199  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
structurally and functionally unrelated lipophilic antitumor drugs.  
C:Genetics:  
A:Gene: mdr1 (pgp1)  
A:Introns: 21/2; 38/3; 92/1; 111/2; 175/2; 233/3; 274/2; 332/3; 370/3; 407/3; 449/3;  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane  
F:1-637/859-1276/Region: duplication  
F:409-603/Domain: ATP-binding cassette homology <ABC1>  
F:426-433/Region: nucleotide-binding motif A (P-loop)  
F:550-554/Region: nucleotide-binding motif B  
F:1051-1247/Domain: ATP-binding cassette homology <ABC2>  
F:1068-1075/Region: nucleotide-binding motif A (P-loop)  
F:1194-1198/Region: nucleotide-binding motif B  
F:73,91,96,103/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:432/Binding site: ATP (Lys) #status predicted  
F:1074/Binding site: ATP (Lys) #status predicted

Query Match 25.3%; Score 1630.5; DB 1; Length 1276;  
Best Local Similarity 31.6%; Pred. No. 1-2e-108;  
Matches 422; Conservative 250; Mismatches 517; Indels 145; Gaps 22;

Qy 3 DESKPRNODGSESSHPPEKETEGRISDYLRIFRYADKYDWTNLVIALICAGSGAS 62  
Db 4 EENLGRADNFKSKMKCKKKEKPAVGFGFMYADWLKMLGTLAAIHGTL 63  
Qy 63 LPLMSIIFSGFTNKN-----NYSNGDGP-----EAFKADVDHFVLMFVLTG 107  
Db 64 LPLMLVFGNMTDSFKABASILPSITNQGPNSTLIISNSLEEMAIYAYITGIGAG 123  
Qy 108 KFLVTYVSTAATPISATIRTRLRVFLLECTLRQVHFKQKNGALATXVTNNGNRQT 167  
Db 124 VLIYAVIOVSLWCLAAGROIHKIQKFFHAINMOEIGFVDVHVDVGLNLTALTDVSKIND 183  
Qy 168 GIAEKLVTFTQALSMFFSAFVALASOKKALITMSVIPALFELVTGICTAIDAAEART 227  
Db 184 GIGDKIGMFFQSITTFPLAGFIIGFSGKLLVILAVSPILGLSSALWAKVLTSTNREL 243  
Qy 228 RIYSRAAVLAEEVLSSIRTVHAFYAQKKMVEKYDVFLOQAQHQEGKKSPNGVLTSTEF 287  
Db 244 QAYAKAGAAVEVLAARTVIAFGGQKLEERYKNLEAKNVGIRKATATASISIGIAYL 303  
Qy 288 CIYAAIALAFWKGFRMYQNGEVADYGVFTVTVLSVTLAATSISMLAPSGSVYVQRIFGS 347  
Db 304 LVYASYALAPWYGTSLVLSNEYS-IGEVLTVFVTSILLCTFSIGHLAPNIEAFANARGAA 362  
Qy 348 ELFSIIDKPTQLDPLDPGKQPEGCGIQEITONLAFAYPSRPSAQVLRDFNLITIPAGT 407  
Db 363 EIFKIIDNPSIDSFSTKGYKPDSTMGNLKFNHFNPSRSEVQILGNLKVKSQTV 422  
Qy 408 ALVCASSGSGKSTVGLIERWYLPSSGRILLDGLGOYNNVWLSRLRLVQOEPVLRGT 467  
Db 423 ALVNSCGKSTTVQLMQRLDPLLEGVYSIDQDIFRINRYLREILGVVSQEPVLFAT 482  
Qy 468 IFQNIANGFMDEQRDLPREKQMLVQKACKASNGDVFINELPNGYTEVEGERAGALSGQ 527  
Db 483 IAEINRYGRED-----VTWDETEKAVKANAYDITMKLPHQFDTPLGEGKAGLSGQ 534  
Qy 528 RORTAARSTISDPKILLDEATSALDPKAEKVQAEALNRVSKDRITVTAHKLATVSKA 587  
Db 535 KORTAARALVRNPKILLDEATSALDTESEAVVQAALDKAREGRTTIVTAHRLSTVRN 594  
Qy 598 GNIAVISQGIKVEGQTHHEIEFGCHYALVRAODLGADEQ-----OPHEKTLHEKAA 640  
Db 595 DVIAFGGIVVQGNDELREKGIYFKLVTQTRGNETEPGNNGAYGSDTDASELTS 654  
Qy 641 REAAGERPALERTHTTATSQAGDLKRRKVPVGTGLYSLLKCI-----LIMFYEQKNL---- 692

Db 655 EES--KSLIRRSIYRSHRQDQERR-----LSMKEAVDEVDVPLVSWRLNLNS 704  
 QY 693 YWCFL--STIVLCAATFGQALLSRLLTVFELSG--HAAQERADYFVILMEFWALGN 749  
 Db 705 ENPYLLVGLCAVINGCIQVFAIVSRVGVFSRDDHETKRONCNLFSLFFVLMGLIS 764  
 QY 750 LVGFTTIGTNCWISQVVTURYQAAMFORVLODDIELDIPQISGALTSLSALPTQLQ 809  
 Db 765 FVTFYFGFTFGKAGEILTARVYMYFKSMLRQDISWFDHKNSTGSLTTLASDASSVK 824  
 QY 810 ELISANFLIYIVGQHRLEQCTSLWNETGPGCVCMCTSTPAFGW-----855  
 Db 825 GAMGARLAV-----VTQNVANLGTGVIL-----SLVYGWOLTVLLVLIPLIV 867  
 QY 856 -----LPQNSRDEARSRLKGLCRKWCACKRSSYRDPDLIFDSRPPCSPTVLGH 906  
 Db 868 LGGIIEMLKLSQALKDKKQLEISGKIATEA-----IENFRIVSLTREOK 913  
 QY 907 VEOGLAKTIO-----SFWEGRGCGFHLQSMEFLAIALGFCIAVDNMLVSTTQLN 957  
 Db 914 FETHYAQSLQVYPYRNAMKKARVGTIPSF--TOAMMFYSYACFRFGAYLVQAQMLTFEN 971  
 QY 958 FISSWAFCLPVQAAQYLAYSTFTKARSAANYLWL-----RTLKPTIRET 1005  
 Db 972 VMLVFSAYVFGAAGNTSFPADYAKAKVSASHIIRIEKTPETDSYSTGLKPTLLEG 1031  
 QY 1006 EENKKGPVGGCPVDLIEDFRYRORDSARVLRGVSMITPQVYVYVAGCGCKSTLIA 1065  
 Db 1032 N-----VKENGVOFNPTPRNPVLPVQLGLSLEVKKGOTLALVGGSGCKSTVQ 1079  
 QY 1066 LSERFYDPTSGRISFAHENTAEPSRLYGRHMSLYOQEPFLVQGSVRENVTLA--LEAEL 1123  
 Db 1080 LLERYDPMAGSVFLDQKEIKOLNVQWLRHLGIYQSEPLFDCSIAENIAYGNSRAVS 1139  
 QY 1124 SEELCOGRALPARMLNTLSLYOKALKRLAORGHQFSGGQQRQIARALIRPKLILL 1183  
 Db 1140 HEEIVRAAKEAN--THQIDSLDPKYNTRV--GDKGTQSGGQQRQIARALIRPKLILL 1197  
 QY 1184 DEATSALDQSERLVQALDEASTRTTIAVAHRLSTIRNVDVFPVFANGRIAGTGHAE 1243  
 Db 1198 DEATSALDTESEKVVQBALDKAREGTCIVIAHRLSTIONADLVIVVIENGKVGKHTHQ 1257  
 QY 1244 LQRLGRYEMCLA 1257  
 Db 1258 LLAQKGIYFSMVAQ 1271

RESULT 8  
 A34786  
 multidrug resistance protein la - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 02-Feb-2001  
 C:Accession: A34786; A35671  
 R:Devault, A.; Gros, P.  
 Mol. Cell. Biol. 10, 1652-1663, 1990  
 A:Title: Two members of the mouse mdr gene family confer multidrug resistance with over  
 A:Reference number: A34786; MUID:90205845  
 A:Accession: A34786  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1276 <DEV>  
 A:Cross-references: GB:M30697; NID:g199111; PIDN:AAA39517.1; PID:g387429  
 R:Hsu, S.I.H.; Cohen, D.; Kirschner, L.S.; Lothstein, L.; Hartstein, M.; Horwitz, S.B.  
 Mol. Cell. Biol. 10, 3596-3606, 1990  
 A:Title: Structural analysis of the mouse mdr1a (P-glycoprotein) promoter reveals the ba  
 A:Reference number: A35671; MUID:90287150  
 A:Accession: A35671  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-525; 'HV', 528-938, 'S', 940-1035, 'V', 1037-1276 <HSU>  
 A:Cross-references: GB:M33581; NID:g199104; PIDN:AAA39514.1; PID:g387427  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P  
 F:406-600/Domain: ATP-binding cassette homology <ABC1>  
 F:423-431/Region: nucleotide-binding motif A (P-loop)  
 F:547-551/Region: nucleotide-binding motif B  
 F:1049-1245/Domain: ATP-binding cassette homology <ABC2>  
 F:1066-1074/Region: nucleotide-binding motif A (P-loop)  
 F:1132-1196/Region: nucleotide-binding motif B  
 F:425/Binding site: ATP (Lys) #status predicted  
 F:1072/Binding site: ATP (Lys) #status predicted

Query Match 25.2%; Score 1627.5; DB 2; Length 1276;  
 Best Local Similarity 32.2%; Pred. No. 1.9e-108;  
 Matches 419; Conservative 244; Mismatches 559; Indels 81; Gaps 18;

QY 3 DESKPPNDGSGESSHPPPEKTEGSDVLRIFRYADKYDWTNLNVIALICAGSGAS 62  
 Db 4 EEDLKGRADNFSKMKCKKKEKKPAVSVLTMPRYAGMLDRMLYLVGTLLAALIHGVA 63  
 QY 63 LPLMSIIFGSEFTNKF-----NNYNSGDSPEAFKADVOHFVLMFVYLF----IGKFV 110  
 Db 64 LPLMLLIFGMDTDSFASGVNSKNTNMSEADKRAMFKLEEMTYAYITGIGAGVLI 123  
 QY 111 LTVYSTAAITISARTTRTLRRRVFLECTLRQEVHPDKOSNGAIAATXVTTNGNRIOGIA 170  
 Db 124 VAYIQVFWCLAAAGROTHKIRQFFHAINNQBGIMFQHDVHGELNTRLTDVDSKINEGIG 183  
 QY 171 EKLVTVOALSMFSAFVALASQWKLALITMSVIPALFVLTGICIAIDAAOAEARITRY 230  
 Db 184 DKICMFQAMATFFGFIIGETRGWKLTLVLAISPVLSAGIWAKLILSFTDKELHAY 243  
 QY 231 SRAAVLAEEVLSSRTVTHAFYAKKMKVRYDVLQQAQHEGCKKSPNNGVLFSTYEYCIY 290  
 Db 244 AKAGAVAEVLAIRTVIAPGQKKELERYNNLEAKRLGKIKKAITANISWGAFLIY 303  
 QY 291 AAIATAFATGFRMYQCEVADGVKVTYVLSVTLAATSLMSLAPSGSVVYORRIFGSELP 350  
 Db 304 ASYALALWYTSVLVSKESYS--IGQVLTVFSLIGAFSVQGNPNIAPANARAAVEVF 362  
 QY 351 SIIDKPTOLDPLDPGKQPEGCLGOIEIONLAPYSPRSQAQVLRDFNLTPACKTALY 410  
 Db 363 KIIDNKPISDFSKGKKPONIOGNLEFNHPSYPSRKEVQILKGLNLKLYKSQOTVALY 422  
 QY 411 GASGSGKSTVMGELLERWYLPSSGRILLDGLGQYNVWKLASRLVQOEPVLRFGTIFQ 470  
 Db 423 GNSGCKSTTVOLMQRLYDPLDGMVSDIGQDITINVRYLREIIGVWSQEDVLPATTAE 482  
 QY 471 NIANGFDEQDRLPREKOMELVOKACKASNGDVFINELPNGVYETEVGERACALSGGQQR 530  
 Db 483 NRYGRED-----VTWDEIEKAVKEANAYDFIMKLPHQFDTLVGERGAQLSGGQQR 534  
 QY 531 IAIARSLIDPKLILLDEATSLADPKAKVQOELNRYSKORTTLVIAHKLATVKSAGNI 590  
 Db 535 IAIARALVRNPKILLDEATSLADSEAVQALDKAREGRTTIVIAHRLSTVYRNADVI 594  
 QY 591 AVISQKIVFQGTGTHHELIFEGCHYAALVRAQDLGADQEOEHE----KTLHEKAAREAGE 646  
 Db 595 AGFDGGVIVFQGHDELREKGIYFFLVMTOTAGNEIELGHEACKSKDEIDNLDMSKDS 654  
 QY 647 RPALERTHTTATSOAG--DLEKRRKVPVGTLYSLKLCILIMFYOKNLVW--CFLLSITV 703  
 Db 655 GSSLIRRRSTRKSTCGPHDQDKLSTKEALDEDVPPASFWRLLKLNSTWPFVVGIFCA 714  
 QY 704 LICAATFPGGALLFSLRLTVFELSG--HAAQERADYFVILMEFWALGNLGYFTIGWTCN 761  
 Db 715 IINGLOPAFVSIVPSKVGVFTNGGPPETORONSRLFSLLLILGILSIFTFPLQGTG 774  
 QY 762 VISOVVTHRQAAMFORVLDQDIELDIPQISGALTSLS-----SALPTQLQ 809  
 Db 775 KAGELTKRLRYVMVFKSMLRQDQVSMFDDPKNTTGALTTLANDAAQVKGATGSLAVIFQ 834  
 QY 810 EL--ISANFLIYIVGQHRLEQCTSLWNETGPGCVCMCTSTPAFGW-----867

Db 835 NIANLGTGIIISLIVG-----WQTLTLLLAIVP--IIAJGVVEMKMLSGQALKDKREL 886  
QY 868 RKLKILKRCWACKRSSYRDPDLRFIFOSRRPCSPYVLGHVVEOGLAKIIQ-----SF 918  
Db 887 EGSGRKATEA-----IENFRTVSLTREKQFETWYQSLQIPIRYNAMKKAH 932  
QY 919 WFCRCFGPHLSOSMEFLAIGFCIAVDNKNVSTTQNLFISSSWAFCLPVQAAQYLAY 978  
Db 933 VFGITIFF--TQMMYFYACFREGAYLYTOQLMTENVLLVFSALVFGAMVGVSSF 990  
QY 979 STSFTKARSAANYIWLKTLKPTITRETEENKKGPGVCGPVDLEDIEFRYRQDRSARVLR 1038  
Db 991 APDYAKATVYASHHIIRIEKTPEDSDYSTQGLKNMLEGVNFQSGFVFNYPTRPSIPVLQ 1050  
QY 1039 GYSMTPEQFVAYVAGSGCKSTLIALSERFYDPTSGRISFAHENIAEMSPRLYRGHMS 1098  
Db 1051 GLSLEVKKGOTLALVGGSCCKSTVQVLLERFYDPMAGSVFLDGKEIKQLNVQWLRAGLG 1110  
QY 1099 LVQOEPFLYOGSVRENTVLAEL--SEELCQGRLPARPMILSSLYQKALKRLAAOR 1156  
Db 1111 IVSQEPFLDCSTAEINAYGNSRVSYEEIVRAAKKAN--IHQFIDSLDPKYNTRY-GDK 1168  
QY 1157 GMQFSGGQRRIATARILIRPKILLDLDEATSDALDQSERLVQAALDEASTRTTAVAH 1216  
Db 1169 GTQLSGGQRRIATARILVROPHILLDLDEATSDALDTESEKVVQOALDKAREGRTCIIVAH 1228  
QY 1217 RLSTRNVVDIVFVFGAETGTHAELQRLRGRYEMCLAQS 1259  
Db 1229 RLSTIQNALDIVVQNGKVKKEGTHQQLAOKGIYFSMVSVQA 1271  
RESULT 9  
DWHYC  
multidrug resistance protein 1 - Chinese hamster  
N:Alternate names: p-glycoprotein pgpl  
C:Species: Crictetus griseus (Chinese hamster)  
C>Date: 31-Dec-1990 #sequence,revision 30-Sep-1992 #text,change 19-Jan-2001  
C:Accession: A38696; C38696; A27126; S33768; 152823  
R:Devine, S.E.; Hussain, A.; Davide, J.P.; Melera, P.W.  
J. Biol. Chem. 266, 4545-4555, 1991  
A:Title: Full length and alternatively spliced pgp-1 transcripts in multidrug-resistant  
A:Reference number: A38696; MUID:91154265  
A:Accession: A38696  
A:Molecule type: mRNA  
A:Residues: 1-1276 <DRV>  
A:Cross-references: GB:M59253; NID:g191154; PIDN:AAA37004.1; PID:g191155  
A:Accession: C38696  
A:Molecule type: mRNA  
A:Residues: 108-1276 <DEL>  
A:Cross-references: GB:M59254; NID:g191156; PIDN:AAA37005.1; PID:g191157  
A:Experimental source: clone ADX185  
A:Accession: B38696  
A:Molecule type: mRNA  
A:Residues: 1-32,771-1276 <DE2>  
A:Cross-references: GB:M59252; NID:g191152; PIDN:AAA37003.1; PID:g191153  
A:Experimental source: clone ADX124  
R:Endicott, J.A.; Juranka, P.F.; Sarangi, F.; Gerlach, J.H.; Deuchars, K.L.; Ling, V.  
Mol. Cell. Biol. 7, 4075-4081, 1987  
A:Title: Simultaneous expression of two p-glycoprotein genes in drug-sensitive Chinese H  
A:Reference number: A27126; MUID:88122132  
A:Accession: A27126  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 706-1276 <END>  
A:Cross-references: GB:M17897; NID:g191158; PIDN:AAA37006.1; PID:g191159  
R:Zastawny, R.L.; Ling, V.  
Biochim. Biophys. Acta 1173, 303-313, 1993  
A:Title: Structural and functional analysis of 5' flanking and intron 1 sequences of the  
A:Reference number: S33768; MUID:93305724  
A:Accession: S33768  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-21 <ZAS>

A:Cross-references: EMBL:L03286  
R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.  
Cell Growth Differ. 2, 425-437, 1991  
A:Title: Analysis of the Chinese hamster p-glycoprotein/multidrug resistance gene pgp  
A:Reference number: 152823; MUID:92088970  
A:Accession: 152823  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-21 <RES>  
A:Cross-references: GB:S81975; NID:g240862  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
tructurally and functionally unrelated lipophilic antitumor drugs.  
C:Genetics:  
A:Gene: pgpl  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding;  
F:407-601/Domain: ATP-binding cassette homology <ABC1>  
F:424-431/Region: nucleotide-binding motif A (p-loop)  
F:548-552/Region: nucleotide-binding motif B  
F:1050-1246/Domain: ATP-binding cassette homology <ABC2>  
F:1067-1074/Region: nucleotide-binding motif A (p-loop)  
F:1193-1197/Region: nucleotide-binding motif B  
F:1193-1197/Region: nucleotide-binding motif B  
F:87,91,96/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:430/Binding site: ATP (Lys) #status predicted  
F:1073/Binding site: ATP (Lys) #status predicted  
Query Match 25.0%; Score 1611; DB 1; Length 1276;  
Best Local Similarity 32.1%; Pred. No. 2.9e-107; Indels 88; Gaps 18;  
Matches 413; Conservative 251; Mismatches 534;  
QY 24 EKTEGSIISDYLRIFRYADKYDWTNLNIALICATGSGASPLMSIIFGSTNKF----- 77  
Db 25 EKKEKPVSVSVTFMYAGWLDRLYLVTAAIIHGVALPLMLVFGDMTDSFASVGN 84  
QY 78 -----NNYNSGGSGSPKADVDHFWLWFLVFLTGKFLV--TYVSTAATITSAIRTRTL 130  
Db 85 PTNATNATQVNASDIFGKLEEMTYAYVYTGIGAGVLIVAYIQVSWFCLAAAGRIHI 144  
QY 131 RRVFLECTLRQEVHFDKOSKAGIAITXTTNGNRITQIGIAEKLVFTVOALSMFSPVVA 190  
Db 145 RQKFFHAINMQNEIGDFVDHVGELNTRTDDVSKINEGIDKIGMFFQAMATFFGGFIIG 204  
QY 191 LASQKALITMSVIPALFVTGICIAIDAAQAEARITRIYSRAAVLAEEVLSIRTVHAF 250  
Db 205 FTRGKWLITVILAIASVPLGSLAGIWAIISSFTOKELQAYAKAGAAVEVLAARTVIAF 264  
QY 251 YAOCKMVKYDVFLOQAHOEGKKSPNNGVLFSTFPCYIAAIALAFWKQPRMYQGEVA 310  
Db 265 GGQKELERYNNLEAKRLGIIKKAITANISMGAAFLIIYASALAFWYGTSLVISKEYS 324  
QY 311 DVCKVFTVLSVTLAATSIISMLAPSGSVYQRRIFGSELSIIDKPTQLDPLDPSKQPE 370  
Db 325 -IGOVITVTFVAVLIAPFSYGOASPNIEAFANARGAAVEIFNIDNKPSIDFSKNGYKPD 383  
QY 371 GCLGQTEIONLAFAPRPSAQVLRDNLITIPAGKTALVAGSGKSTWVGLERWYLP 430  
Db 384 NIKGNLEFNHIFSPSRKDVQILKGLKVQSGTVALVGNVSGCKSTTVQLQLRLYDP 443  
QY 431 SSGRIILLGELGVQNVKWLRSFRLVQOEPVLPFRGTIFONTIANGFEMDEQDLPREK-OM 489  
Db 444 TEGVVSIDQDITINRYLREIIIGVVSQEPVLPFATIAENIRYG-----RENVTV 494  
QY 490 ELVOKACKASNGDVFINELPNGYETEVGERAGALSGGORIATARSIIIDPKILLDEA 549  
Db 495 DEIEKAVKANAYDTIMKLPHKFDPLVGERGAQLSGGQKQRIATARALVRNPKILLDEA 554  
QY 550 TSALDPKAKVQVEALNRVSKDRTLVIAHKLATVKSAGNIATVISOQKIVEOGTHHELT 609  
Db 555 TSALDTESEAVVQAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGCVIYVEQGNHELMR 614  
QY 610 FQCHYAALVRAODIGAD-----EQQEHKTLHEKAREAMGERPALETHTTATSGAT 652

Db 615 EKGIYFKLVMTQTAGNIEELNCEVGESEKNEITDNLDMSSKDSAS---SLIRRRSTRSIRG 671  
QY 663 --DLEKRVKVPVCTGLYSLLKILLIMFYEOKNLYW-CELLSTIVILCAATFGCOALLFSR 719  
Db 672 PHQDRKLISTREALDEDVPPISFRILKLSNSEMPYFVGIFCAIVNGALQAPAFSIFSK 731  
QY 720 LITVVELSCHAAERAD--FYILMFVVALGNLVGYFTIGWTCNVISQVVTTHRYQAMFP 777  
Db 732 VVGVTFRNTDDETRKHSNLSLLFLIIGVLSFTTFPLOGTECKAGEILTKRLRYMVKF 791  
QY 778 RVLDQDIELLDPEIGISALQSLS-----ALPTQLBELSANFLIYIVVG 823  
Db 792 SMLRDVSNFQNPKNYTCALTFRLANDAGOVKGATGARLAVITQNTIANLGTGIIISLI 851  
QY 824 QHRLSOCSTTSLMWTGPGCVWCTSTPAFCWLPONSSRDARSRLKGLRCKWACKRS 883  
Db 852 -----WQLTLLALLAIVP--IIAIVGVEMKMLSGOALKDKKLEGGSKIAIEA----- 897  
QY 884 SYRDPRLIFOSRRPCSTVLGHVDEGLAKIIQ-----SFWFGRCFGFHLQSMEF 934  
Db 898 -----IENFTVVSILTRQKPFENMYAQSLOIPYRNALKKAHVFGITFSF--TOAMMY 947  
QY 935 LAIALGFCIAVDNWLGVSTTQNLNFISSSWAFCLPVOAAQYLAYSTSTFKARSAANYILM 994  
Db 948 FSVACRFEGAVLVARLMTENLVLSAIVFCAMAVGVSSFAFDYAKAKVSAASHIIM 1007  
QY 995 LRLTPTTRETENKKGPGVGCPOVLEDETFYRQRDSARVLGRVSMTEIPGQFVAYVG 1054  
Db 1008 IIEKVPISDSYSTGGLKPTNLEGVNFVFNVPYTRPDIPVLQGLNLEVKGGOTLALVG 1067  
QY 1055 ASGCKSTLIALSERFYDPTSGRISFAHENTIAEMSPRLYRGHMSLVQOEPYLYQGSVREN 1114  
Db 1068 SSGCKSTVOLLERYFDPMACTVFLDGKEVNLQWLNQWLAHGLIVSQSPILFDCSIAEN 1127  
QY 1115 VTLAEAE--LSEELCOGLPARPMLWILSSLYOKALKRLAAOQMGFSGGQORATARA 1173  
Db 1128 IAYGNSRVVQDEIERAKCANIHQIESLPDKYNTRY--GDKGTQLSGGQKQRIARA 1186  
QY 1174 LRNPKLLLDATSDTOSERLVQALDEASTRTTIAVAHRLSTIRNVDVVFVANG 1233  
Db 1187 LVHQPHILLDEATSDTSEKVVQBALDKAREGTTCIVIAHRLSTIONADLVIVQNG 1246  
QY 1234 RIAETGTHAELQRLGRYEMCLAQS 1259  
Db 1247 VKYKHGTHQQLAOKGIYFSMVQOA 1272

## RESULT 10

## DMS2

multidrug resistance protein 2 - mouse  
N:Alternate names: P-glycoprotein MDR2  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
C:Accession: A30409; S70711  
R:Gros, P.; Raymond, M.; Bell, J.; Housman, D.  
Mol. Cell. Biol. 8, 2770-2778, 1988  
A:Title: Cloning and characterization of a second member of the mouse mdr gene family.  
A:Reference number: A30409; MUID:88302195

A:Molecule type: mRNA  
A:Residues: 1-1276 <HSU>  
R:Cross-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428  
R:Kirschner, L.S.  
Nucleic Acids Res. 24, 2829-2834, 1996  
A:Title: De novo generation of simple sequence during gene amplification.  
A:Reference number: S70711; MUID:96313253  
A:Accession: S70711  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 43-92 <KIR>  
R:Cross-references: EMBL:U46839; NID:g1228142; PIDN:AAC52722.1; PID:g1228143  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell

structurally and functionally unrelated lipophilic antitumor drugs.

C:Genetics: mdr2

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembr  
F:1-637,653-1276/Region: duplication  
F:409-603/Domain: ATP-binding cassette homology <ABC1>  
F:436-433/Region: nucleotide-binding motif A (P-loop)  
F:550-554/Region: nucleotide-binding motif B  
F:1049-1245/Domain: ATP-binding cassette homology <ABC2>  
F:1067-1074/Region: nucleotide-binding motif A (P-loop)  
F:1192-1196/Region: nucleotide-binding motif B  
F:98-94/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:432/Binding site: ATP (lys) #status predicted  
F:1072/Binding site: ATP (lys) #status predicted

Query Match 24.8%; Score 1601.5; DB 1; Length 1276;  
Best Local Similarity 32.4%; Pred. No. 1.4e-106;  
Matches 423; Conservative 243; Mismatches 523; Indels 115; Gaps 26;

QY 17 SSSHPPEKETEGSISDYLRIFRYADKYDWTUNVIALICAGSGASLPLMSIIFGSPFNK 76  
Db 22 SINSQREKKKKVNLGLLTFLRYSDWQDKLPMFLGTLMAHAGSLPLMIVFEGMTDK 81  
QY 77 FNNYNSGDGS-DEAF-----KADVHFVLFVYFLIGKFPVLYVSTAITSAT 124  
Db 82 FVD-NTGNFSLPNVFNLSMLNFCRILEEMTRYAYYYGLGGVLYAAIYQVSWTAAAG 140  
QY 125 RTTRTLLRRVFELECTLRQEVYHFDKQSGAIAATXVTNGNRIQTGTAEKLVTVQALSMFF 184  
Db 141 RQKKIRQKFFHAILROEMGFEDIKGTTELNRLTDVDSKISEGIDKRVGMFFQAIATFF 200  
QY 185 SAFVVALASQWKLALMTSVIPAILVPTGICIAIDAQAEARITRYSAALVALEVLSSI 244  
Db 201 AGVIFVIRGKLTIVIMAIISPILGSTAVWAKILSTFSDKELAAKAGAVAEAPGAI 260  
QY 245 RYHAYVAKKVEKYDVPLOQAHOEGKKKSPNNGVLFSTFYFCIYIAIALAFWKGRMY 304  
Db 261 RTVIAFGGQNKELERYOKHLENKXIGIKKAIASINSGIAFLIYASIALAFWYGTSLV 320  
QY 305 QNGEADVAGKVTVLSVTLAATISISMLAPSGSVVYQRRIFGSELSFISIDPTOLDLDP 364  
Db 321 ISKEYT-IGNAMTVFESIUGAFSVGQAAPCIDAFANARGAAYVIFDIIDNNPKIDSFE 379  
QY 365 SGKQPSGLQTEIONLAFAYSPRSQAVLRDPLNTIPAGKTTALVAGSGSKSTWGLL 424  
Db 380 RGHKPNIGNLEFSDVHFYSRANIKLGLNLKVKSGQTVLVGNVSGCKSTTVOLL 439  
QY 425 ERWYLPSSGRILLDGLGQYVNWKLRSRIRLVQOEPVLPFGTIPQNT----ANGFMDQ 480  
Db 440 QRLYDPTGKISTDQDRIENFNRCLREIIGVVSQEPVLPFTTIAENIRYGRGNTWDE- 498  
QY 481 ROLPREKOMELVQKACKANGSDVFINELPNGVETEVGERAGALSGGQORAIARSIISD 540  
Db 499 -----IEKAVKEANAVDFIMKLQKFDTLVGDGQALSGGQKORAIARALVRN 547  
QY 541 PKTLLDEATSDALDKPAKVEQAEALNRVSKORTTLVIAHKLATVKSAGNIAVISOGKIVE 600  
Db 548 PKTLLDEATSDALDTESEAEVQALDKAREGRTTIVIAHRLSTIRNADVIAGFEDGVIVE 607  
QY 601 QGTHHELIFGCHYAAALVRAQDLGAD-EQQEHEKTL-HEKAAREAA--GERPALERTHTT 656  
Db 608 QGSHSELMKKEGIYPRLVNMTAGSQILLSEFEVELSDEKAAGDVAPNGWKARIFRNSTK 667  
QY 657 AT-----SQAGDEKRRKVPVGTGLYSLKLCILLMEYEQKNLYWCFLSLTITVLIC 706  
Db 668 KSLKSPHONRLDEETNELDANVPPV-----SFLKVLKL----NKTEWPFVVGTVCAIAN 718  
QY 707 AATFPQOALLFRLTLVTFELSGHAA-QERADFYILMEFVVALGNLVGYFTIGWTCNVISQ 765  
Db 719 GALQPAFSLISEMIAIFGPGDDQVQKQKNMFSLVFLGLGLVFSFTFFLQOGFTFGKAGE 778

QY	766	VYTHRYQAAWFORYLDQDIELDDIPQISGALTSQLSALPTQLQ-----ELISAN----	815
Db	779	IIITLRLSMFAKMLRQMSWFDHDKNSTGALSTRLATDAAOVGATGKTALIAQTAN	838
QY	816	-----FLIYVVGQHRLEQCSITSLWMTGP-----GCVWCTSTPAQWLPONSSRDERS	867
Db	839	LGTGIIIIISFTYG-----WQLTLLLSLVYPPIAVAGIV-----EMKLAGNAKRDKKEM	886
QY	868	RKLGLKRLCKWACKRSSYRDPDLRIIFDSRRSPCSPTVLGHVQEGLAKLIQSEWFCRCFGEH	927
Db	887	EAAGKATEAETINRTVYSLVTOERKFESM-----YVEKLHGPYNSVRKAHIYGIT	937
QY	928	LSQWMEFLATALGCIADVNNMLQVSTQLNF-----ISSWAFCLPVOAAQAQLAYSTSEFK	984
Db	938	FSISOAFMYESYAGCPRFGSYLIYN-GHMRKQVILVFSAILVGLAVALGHASSFAPDYAK	996
QY	995	ARSAANYILMLRTLTKTIRETE-----ENKKKGPVGCPOVLEDIEFPYRORDSARVLG	1039
Db	997	AKLSAAVILFSLPRQPLIDISVSGCLMPDFEG-----SVTFNVFVNPTRANVPVLQG	1051
QY	1040	VSMTRIGQVAVVVGASGGCKSTLIASERFYOPTSGRTSFAHENIAEMSPRLYRGHMSL	1099
Db	1052	LSLGVKAGQATLALVSGSGCKSTVQLLERYPYDPMAGSVLLDQBAKLNQWMLRAQLGI	1111
QY	1100	VQOEPFTLYOGSVRENTYLTALAEAL--SEELCOGRLEPA--RPMLWLITSLSYOKALKRLAAQ	1155
Db	1112	VSQSPILFDSCIAENIAYGDSRVVPVPHDEIVRAKEANIHPETLPQKYNTRV-----GD	1167
QY	1156	RGMOFSGGQRIQRIATARILIRPKLLLLDEATSAIDTQSERIAVOALDEASTSTRTTAVA	1215
Db	1168	KGTOLSGGQRIQRIATARILIRQPRVLLLDEATSAIDTQSEKVVQVQALDKAREGRTFVIA	1227
QY	1216	HLRSTINVDIVFPANGRIAGTGHAEQLQRLGRYEMCLAQS	1259
Db	1228	HLRSTIQNALIVVIENGKVRKHGTHQQLLAQKGIYFSMVNIQA	1271
RESULT 11			
T19982			
Hypothetical protein C47A10.1 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000			
C:Accession: T19982; T24202			
R: Basham, V.			
submitted to the EMBL Data Library, November 1996			
A:Reference number: Z19207			
A:Accession: T19982			
A:Status: preliminary; translated from GB/EMBL/DBDJ			
A:Molecule type: DNA			
A:Residues: 1-1294 <W1>			
A:Cross-references: EMBL:Z81484; PIDN:CA803973.1; GSPDB:GN00023; CESP:C47A10.1			
A:Experimental source: clone C47A10			
R: Basham, V.			
submitted to the EMBL Data Library, April 1997			
A:Reference number: Z19853			
A:Accession: T24202			
A:Status: preliminary; translated from GB/EMBL/DBDJ			
A:Molecule type: DNA			
A:Residues: 1-1294 <W1>			
A:Cross-references: EMBL:Z93782; PIDN:CA807855.1; GSPDB:GN00023; CESP:C47A10.1			
A:Experimental source: clone R1268			
C: Genetics:			
A:Gene: CESP:C47A10.1			
A:Map position: 5			
A:Introns: 15/2; 67/3; 123/3; 168/2; 509/1; 649/3; 826/2; 1089/3; 1158/3; 1250/3			
C: Superfamily: multiaug resistance protein; ATP-binding cassette homology			
Query Match 24.8%; Score 1601.5; DB 2; Length 1294;			
Best Local Similarity 32.1%; Pred. No. 1.4e-106;			
Matches 433; Conservative 234; Mismatches 509; Indels 175; Gaps 333			
QY	7	KPRPNQDCSESSSH-----PPPEKTEGTSIDYLRIFRYADKYDWTLVNIALICAGSGA	61





C:Accession: S41646; S22354  
 R:Brown, P.C.; Thorgeirsson, S.S.; Silverman, J.A.  
 Nucleic Acids Res. 21, 3885-3891, 1993  
 A:Title: Cloning and regulation of the rat mdr2 gene.  
 A:Reference number: S41646; MUID:93376516  
 A:Accession: S41646  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1278 <BRO>  
 A:Cross-references: EMBL:U15079; NID:g310192; PIDN:AA02937.1; PID:g310193  
 B:Deuchars, K.L.; Duthie, M.; Ling, V.  
 Biochim. Biophys. Acta 1130, 157-165, 1992  
 A:Title: Identification of distinct P-glycoprotein gene sequences in rat.  
 A:Reference number: S22351; MUID:92223089  
 A:Accession: S22354  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1211-1278 <DEU>  
 A:Cross-references: EMBL:Y61105; NID:q56892; PIDN:CAA43417.1; PID:g1334220  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop  
 F:409-603/Domain: ATP-binding cassette homology <ABC1>  
 F:426-433/Region: nucleotide-binding motif A (P-loop)  
 F:1051-1247/Domain: ATP-binding cassette homology <ABC2>  
 F:1068-1075/Region: nucleotide-binding motif A (P-loop)

Query Match 24.7%; Score 1592.5; DB 2; Length 1278;  
 Best Local Similarity 32.4%; Pred. No. 6.3e-106;  
 Matches 422; Conservative 243; Mismatches 529; Indels 107; Gaps 25;

QY 17 SSSHPPPKTEGSDYLRIFRYADKYDWTNLNLTALICAGSALPLMSIIFGSFTNK 76  
 DB 22 SINSQREKKKKVNLPLTLERYSDQKLFMLLGTAMATAHSGPLMIVFGEWMDK 81  
 QY 77 FNNYNSGDS-PEAF-----KADVDHFLVFLYFQKGFVTVYTAATISAI 124  
 DB 82 FVD-NAGNFSLPVNFSLMNLGRILEEMTRYAVYVGLGGVLLAAYTQVSWFLAAG 140  
 QY 125 RTTRTRLVFLECTROEVNHFQKQSGAIAATXVTNNGRINTQTGIAEKLVFTVQALSMFP 184  
 DB 141 ROIRKIRQKFHAILRCQMGWFDIKGTTELNLTRLTDIDSKISGIGDKVGMFOAIAITF 200  
 QY 185 SAFVVALASOKLALITMSVIPAFIVTGVICITADAAQEARITRIYSRAVLAEEVLSS 244  
 DB 201 AGFVGFIRGNKLVIVKAITALGLSTAVWAKILSTFSDKELAAAYAKAGAAEALGAI 260  
 QY 245 RTVHAFYAKKWKVDYVFOOAHQEGKKSPNGVLFSTFPCIVYAAIALAFWGFMY 304  
 DB 261 RTVIAPGQNKELRYQKHLENAKKIGIKKASINSMGIAFLIYASALAFWYGSILV 320  
 QY 305 QNGEVADGVKFTVSVTLAITSMLAPSGSVYVQRIFGSELSIIDKPTQLDPLDP 364  
 DB 321 ISKEYT-IGNAMTVFFSILIGAVSGQNAAPCIDAFNARGAAYVIFDIIDNPKIDTSFE 379  
 QY 365 SKQPEGCGLOIETONLAFAYPRPQAQVLRDNFTIPAGKTALVAGSGSGKSTWGLL 424  
 DB 380 RGHKPDSTKGNLEFSDVHFYSYPRANIKILKGLNKKVSGQTVALVGNCGCKSTVQLL 439  
 QY 425 ERWILPSSGRILLDGLGQYNYKWLRSIRLVQOEPVLFRTGIFQNI----ANGFMDRQ 480  
 DB 440 QRLYDPTETISIDQDIRNFNRCLREFIGVVSQEPVLFSTTIAENIRYGRGNVTMDR- 498  
 QY 481 ROLPREKQWELVQKACASGNDVFINELPNGETEVGRACALSGGQRIATARSISD 540  
 DB 499 -----IKKAVKEANAYDFIMKLPOKEDTLVGDGQALSGGQKQRIALARLVN 547  
 QY 541 PKILLDEATSDLPKAEKVVQALNRVSKDRITLVIAHKLATYKSNLAVISOCKIVE 600  
 DB 548 PKILLDEATSDLPKAEKVVQALNRVSKDRITLVIAHKLATYKSNLAVISOCKIVE 607  
 QY 601 QGTHHELFECVAAVLAQDLDGAD-EQOEHEKTL-HEKAAREAA--GSRPALERTHT 656

DB 508 OGSHELIIKEGIYFRLVNMQTSQSQILSEFEVELSDEKAAGVAPNGWKARIFRNSTK 667  
 QY 657 ATSOAG-----DLEKRPVPGVLGYSLKLCILIMFYEQKNLYWCFLLSITVILCAAT 709  
 DB 668 KSLKSSRAHONRLDVTNELDANVPVFLKVLRL-----NKTEWYFVGVLCALANGAL 723  
 QY 710 FPGQALLFSLRLTVFELSGHAA-QERADFYILMFVVALCNLYGVFTTGTWCNVISOVVT 768  
 DB 724 QPAFSIILSEMAIRPGDGTVKQCKNMESLFLGLGVHSEFFTEFLQGFITFGKAGELLT 783  
 QY 769 HRYQAMFORVLDDIETLPEOISGALTQSLSALPTQLQ-----ELISAN----- 815  
 DB 784 TRLSRMAFKAMLRQDMWFDHKNSTGALSTRLATDAQVQATGTRLALTAQJANLGT 843  
 QY 816 -FLIYVVGQHRLEOCSTTSLAMETGP----GGCVWCTSTPAFGWLPQNSRSDARSRLK 870  
 DB 844 GIISFIYG-----WQLTLLLSVVPFIAVAGIV-----ENKMLAGNAKRDKKEMEA 891  
 QY 871 GKLCRCWACKRSYRDPDRLIIDFSDRRPCSPTVLGHVQGLAKIITQSFWFCGCFHLSQ 930  
 DB 892 GKIAATEAENIRTVSLTQERKFESM-----YVEKLHGPYRNSVRKAHIYGITFSI 942  
 QY 931 SMEFLATLGFCLAVDNMLQVSTTOLNF--ISSWAFCLPYQAAAQYLAISTSTFKARS 987  
 DB 943 SOAPMYFSYACGFRGCSYLIYN-GHMRKDVILVFSALVGLGVALGHASSFAPDYAKAKL 1001  
 QY 988 AANYILMLRTLPRTTETEE-----NKKKGPVGCPCVDLETFEYRORDSARVLGVSM 1042  
 DB 1002 SAATLSLFRQLIDTSYREGMPDKFEG-----SVTFENVFNYPTKAVPVVQLGSL 1056  
 QY 1043 TIEPGQVATVYAGSGCKSTIALSERPYDTSRISFAHENIAEMSPRLIRGHMSLVQ 1102  
 DB 1057 EVKRGQTLALVSGSGCKSTVQLLERYEDMAGTVLLDQEAUKLUNQWMLRAQGVISQ 1116  
 QY 1103 EPTLYOGSVRENTVLAEL--SEELCQGRIPA--RPMWLILSSLYQAKRLAAOGRM 1158  
 DB 1117 EPIFLDCSIKNTAYGDNRSVVSODEIVRAAKEANHPFETLPQKYEYV---GDKGT 1172  
 QY 1159 QFSGGQRIATARALIRNPKLILLDEATSDALDQSERLVQALDEASTSRITIAVAHRL 1218  
 DB 1173 QLSGGQRIATARALIRNPVLLDEATSDALDQSERLVQALDEASTSRITIAVAHRL 1232  
 QY 1219 STIRNVDIVFVANGRIATGTHAELQRLGRVYEMCLAQS 1259  
 DB 1233 STIONADLIWIDNGVKKEGTHQQLAQKGIYFSVMVNIQA 1273

RESULT 14

148123  
 P-glycoprotein isoform III - Chinese hamster  
 C:Species: Cricetus griseus (Chinese hamster)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 02-Feb-2001  
 C:Accession: U48123  
 R:Accidcott, J.A.; Sarangi, F.; Ling, V.  
 DNA Seq. 2, 89-101, 1991  
 A:Title: Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene fa  
 A:Reference number: U48121; MUID:92135896  
 A:Accession: U48123  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1281 <RES>  
 A:Cross-references: GB:M60042; NID:g191168; PIDN:AAA68885.1; PID:g191169  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop  
 F:412-606/Domain: ATP-binding cassette homology <ABC1>  
 F:429-436/Region: nucleotide-binding motif A (P-loop)  
 F:1054-1250/Domain: ATP-binding cassette homology <ABC2>  
 F:1071-1078/Region: nucleotide-binding motif A (P-loop)

Query Match 24.6%; Score 1590.5; DB 2; Length 1281;  
 Best Local Similarity 32.4%; Pred. No. 8.8e-106;  
 Matches 421; Conservative 242; Mismatches 531; Indels 107; Gaps 24;





-----

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 07:58:40 ; Search time 24.91 Seconds  
(without alignments)

354.654 Million cell updates/sec

Title: US-09-882-694A-11

Perfect score: 6455  
Sequence: 1 MADESEKPRNDGSESSSH.....LQRLGRYEMCLAQLDQA 1263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 40831 seqs, 6994805 residues

Total number of hits satisfying chosen parameters: 40831

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1640	25.4	1279	1	PCT-US02-07826-2
2	1640	25.4	1279	6	US-10-097-340-2
3	1553.5	24.1	1286	5	US-09-573-655B-1211
4	1495	23.2	1286	5	US-09-573-655B-1295
5	682.5	10.6	612	5	US-09-540-209B-6953
6	546.5	8.5	678	5	US-09-573-655B-1015
7	544.5	8.4	659	5	US-09-540-209B-6188
8	536.5	8.3	1123	5	US-09-573-655B-1014
9	512	7.9	690	6	US-10-102-806-726
10	509	7.9	608	5	US-09-540-209B-6291
11	498	7.7	740	5	US-09-540-209B-8625
12	490	7.6	735	5	US-09-540-209B-5749
13	468	7.3	574	5	US-09-540-209B-6190
14	457	7.1	740	5	US-09-540-209B-8427
15	434	6.7	475	6	US-10-076-157-4
16	384	5.9	215	5	US-09-540-209B-6289
17	365.5	5.7	1050	5	US-09-573-655B-2417
18	355	5.5	352	7	US-60-365-384-516
19	311.5	4.8	555	5	US-09-913-414-4
20	291.5	4.5	1434	5	US-09-573-655B-748
21	262	4.4	412	5	US-09-540-209B-10304
22	241	3.7	260	5	US-09-540-209B-8144
23	238.5	3.7	2261	5	US-09-984-827-132
24	237.5	3.7	258	5	US-09-540-209B-6766
25	237	3.7	242	5	US-09-540-209B-10155
26	237	3.7	267	5	US-09-540-209B-9622

27	236	3.7	147	6	US-10-106-698-6340	Sequence 6340, Ap
28	233.5	3.6	228	5	US-09-540-209B-5496	Sequence 5496, Ap
29	233.5	3.6	2261	5	US-09-984-827-128	Sequence 128, App
30	233.5	3.6	2261	5	US-09-984-827-129	Sequence 129, App
31	233.5	3.6	2261	5	US-09-984-827-131	Sequence 131, App
32	233.5	3.6	2261	5	US-09-984-827-136	Sequence 136, App
33	232.5	3.6	470	5	US-09-540-209B-9925	Sequence 9925, Ap
34	232.5	3.6	2261	5	US-09-984-827-127	Sequence 127, App
35	232.5	3.6	2261	5	US-09-984-827-130	Sequence 130, App
36	232.5	3.6	2261	5	US-09-984-827-133	Sequence 133, App
37	232.5	3.6	2261	5	US-09-984-827-135	Sequence 135, App
38	232.5	3.6	2261	5	US-09-540-209B-5412	Sequence 5412, Ap
39	229.5	3.6	227	5	US-09-984-827-134	Sequence 134, App
40	229.5	3.6	2261	5	US-09-540-209B-8138	Sequence 8138, Ap
41	225	3.5	257	5	US-09-540-209B-5266	Sequence 5266, Ap
42	217	3.4	284	5	US-09-540-209B-9107	Sequence 9107, Ap
43	217	3.4	531	5	US-09-540-209B-9107	Sequence 245, App
44	210.5	3.3	323	6	US-10-107-431-245	Sequence 8196, Ap
45	210	3.3	300	5	US-09-540-209B-8196	

#### ALIGNMENTS

RESULT 1

PCT-US02-07826-2

; Sequence 2, Application PC/TUS0207826

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc. et al.

; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, Assessment, Prevention, and Therapy of Ovarian Cancer

; FILE REFERENCE: Mfr-030PC

; CURRENT APPLICATION NUMBER: PCT/US02/07826

; PRIOR FILING DATE: 2002-03-14

; PRIOR APPLICATION NUMBER: 60/276,025

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 60/325,149

; PRIOR FILING DATE: 2001-09-27

; PRIOR APPLICATION NUMBER: 60/276,026

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 60/324,967

; PRIOR FILING DATE: 2001/09/26

; PRIOR APPLICATION NUMBER: 60/311,732

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: 60/325,102

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/323,580

; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 1279

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US02-07826-2

Query Match 25.4%; Score 1640; DB 1; Length 1279;

Best Local Similarity 32.3%; Pred. No. 4.2e-119;

Matches 428; Conservative 243; Mismatches 489; Indels 164; Gaps 27;

QY 24 EKETSGISDYLRIFRYADRYDWTNLNVALICAGSGASGLPLMSIFGSGFTNFNN---79

DB 27 KKEKPTVSVF-SMFRYSNLDKLYMVVGTAAIITHGAGLPLMLVFGEMTDIFANAGNL 85

QY 80 -----YNSGDSGSEAF---KADVDHFVLFVLTGKFTLVTVSAAATTSIAITTR 128

DB 86 EDLMSNITNRSDINDTFFNNLEEDMTRVAYYYSGIGAGLVAAIYQVSWCLAAGRQIH 145

QY 129 TLRRVFLCTURQVWFHQKNGAATATVTTNGNRIQTGAELKLVFTVQALSMFFSAFV 188

DB 146 KIRKQFFHAIHQEIGWFDVHVGELNRLTDDVSKINEGIDGKMGFMFSNATFFTFGI 205





```
QY 649 ALRTHTTATSONGLDEKRR-----VPGVTGL-----YSLKCLL 683
Db 652 SLGRSLKSGSRGSRHSRHSFNMFGFAGIDGNVVOOEEDD/TOPKTPKVKVSPRIAA 711
QY 684 IMFYEQNLVYGLLSTITVLICATPGQALLSRLLTVFELSGHAAQFRADFYLAFF 743
Db 712 L-----NKPEIPVILGSLGSAANGVILPIFGLISSVKAFFQPKKLRKEDTSFWALIFM 767
QY 744 VVALGNVY-----FPIGTGNCVISOVTHRYOAAAFORVLDODIELLDPEQTSAL 797
Db 768 VLGFASIIAYPAQTEFAIA-GCKLV-----QIRSNCFKVVHMEVGFDEPENSSTGI 821
QY 798 TSOLSAALPQLQELISANF-----LIIYVVOGHRLEOCSTTS 834
Db 822 GARLSADAATIRGLVDSILAQTVONLSLILAGLIIAFACQALFVVLAMLPILALN-GF 880
QY 835 LAMETGCGCVNCTSTPAFGWLPQNSRDRSRLKGLCKRCKWACKRSYRDPORLIPD 894
Db 881 LYMKFMKG--FSADAKWIGASQVANDAVGSIKTVASFCAE-----DKVMNM 926
QY 895 SRPCSTVLGHVEOGLAKIIOSWFGRCFGFHLUSQSMFLALALGFCIA---VDNWLQV 951
Db 927 YSKKCGPMKNGIROGVSGI-----GFGSFVLFSVAASFYVGBRLVDDG--- 974
QY 952 STTQLNFSSWAFCLPQVQAAQYLAYSTSTFKARSAANYILMLRLKPTIRETEENKK- 1010
Db 975 KTFDFVFRVFALTAAMAATISQSSLSPOSSKADVAASIFAIDRESKIDPSVESGRV 1034
QY 1011 ----KGPVGGCPVDLEDIEFRYRDRSARVLRGVSMTEPQFVAVVGASGCGKSTLIAL 1066
Db 1035 LDNVKG-----DIELRHVSFYKPARPDVQIFQDCLISIRAGTKVALVGGSGKSTYVAL 1089
QY 1067 SERFYDPTSGRISFAHENAEMSPRLYRHSILVQOEPTLYQGVRENWTLALAEALSEE 1126
Db 1090 LQRFYDPSGEITLQGVKLSRLKWLKQOTGLVQSPILFNETIRANAYKGGGDAGES 1149
QY 1127 LCOGRLPARPMWLILSSYLKALRLAAQRMQFSGGQRORIAARALINPKLILLDEA 1186
Db 1150 EIVSSAELSNANGFTSGI--QGGYDTVMGREGIOLSGGQRORIAARAIKPKVILLDEA 1208
QY 1187 TSALDQSRVQAALEASTSRTTIAVAHRLSTIRNVDFVIFVANGRIATGTTHAELQR 1246
Db 1209 TSALDAESRVVQDALRVMVNRITIVVAHRLSTIKNADVIAYVKNGVIVKGRKHTLIN 1268
QY 1247 LRGRYY 1252
Db 1269 IKDGVY 1274
```

## RESULT 4

```
US-09-573-655B-1295
; Sequence 1295, Application US/09573655B
; GENERAL INFORMATION:
; APPLICANT: SOLOVYEV, Victor and TROUKHAN, Maxim
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0876P
; CURRENT APPLICATION NUMBER: US/09/573, 655B
; NUMBER OF SEQ ID NOS: 3281
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1295
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-573-655B-1295
```

```
Query Watch
Best Local Similarity 30.8%; Score 1495; DB 5; Length 1286;
Matches 401; Conservative 245; Mismatches 530; Indels 126; Gaps 28;
```

```
QY 13 DGSSESSPPPP-----EKETEGSIDYLRIFRYADKYDWTNLNVIALICAGSGASLPLM 66
Db 2 DNDGGAPPTPLTVVEEPKKAERIGVAFKELFHPADGLDVLVLMGIGSVGAFVHGGSLPLF 61
QY 67 SIIFSGFTNKNFNVNSGDSPEAFKADVDHFLWVYVLFITGKFLVTVVSTAATITSAIRT 126
Db 62 LRFFADLVNSF---GNSNNVKKMEBVLKVALYFLVVGAAIWASSWAEISCMWMSGERQ 118
QY 127 TTTLRVFLLECTLRQVYVHFDKQ-SNGAIATXVTTCNRIQTGIAEKLIVTQVQALSMFEFS 185
Db 119 TTKMRIKVLEAALNQDIOQFDEVRTSDVVFATINDAVVQDAISEKLGNFHYMATFVS 178
QY 186 AFVVALASOKLALITSVIPAIFLVTGICIAIDAAQAEARITVYISRAAVLAEEVLSIR 245
Db 179 GRIVQFTAVQWALVTLVAVPLIAVIGIHTTSLKLSNKSQESLSOAGNIYQIVQVIR 238
QY 246 TVHAFYAKKVKVEYDYFLOOAHQEGKKSPNNGVLFSTEYFCIYAAIALAEKWFGRMYQ 305
Db 239 VYMAFVGSERASQAYSSALKTAQKLGKYGKTLGAKMGGLGATYFVFCYCALLWYGGYLR 298
QY 306 NGEVADVGKFTVVLVSLAATSIISMLAPSGSVYVQRRIFGSELSIIDKPTOLDPLDPS 365
Db 299 H-HLTNGGLAIATMPAVNIGGLALGOSAPSMAAFKAKVAAAKIFRIIDHKPTIERNSES 357
QY 366 GKQPEGCLGQIEIQNLAFAYPSRPSAOVLRFNLTIPAGKTTTALVGASGSGKSTVUGLLE 425
Db 358 GVELDSVTGLVELKNVDFSPSRPDVKILNNFCLSPVAGKTTIALVGSOGSGKSTVVSLE 417
QY 426 RWLPSGGRILLDGLGOVNVKWLRSRIRLVOQEPVLFRTIFONIANGFMDQORDLPR 485
Db 418 RYDPSNGOVLDDQDLKTLKRLWROOIGLVQSOEALFATSIKENILLG-----RPD 470
QY 486 EKQMLVOKKACKSNGDVFINELPNGYETEVGERAGALSQGGQORIIATARSIIIDPKILL 545
Db 471 ADQVE-IEEAAVANANASFIILKPDGDTQVGERGLQSGGQORIIATARALKMPAILL 529
QY 546 LDEATSAIDPKAEKVQOEALNRVSKORTILVIAHKLATVKSAGNTAVISQGVKIVGQTHH 605
Db 530 LDEATSAIDSEKLVQOEALDRFMIGRTTIIIAHRLSTIRKADLVAVLQOQSVSIGTHG 589
QY 606 ELIEFGCH--YAALVRAQDLGADQOEKHTLH--EKAAREAAGERPALERTHTTATSOA 661
Db 590 ELFSKGENGVAKLIKQEA--HETAMSNARKSSARPSSARNSSVSPIMTRNSY 643
QY 662 G-DLEKKKVPVGLVYSL-----LKCILIMEPEQKLYWCF-----LLS 699
Db 644 GRSPYSRRLSDSTSDFSLSIDASSYPNYRNEKLFKQDQANSFWRKAKHNSPENKVALLG 703
QY 700 TITVLICATFPGOALLFSRLTVFELSGHAAQ-BRADFYILMFFVWALGNLVGYETIG- 757
Db 704 SVGSVTCGSLSAFPAYVLSAVLSVYNPOREYMIKQIDKXCYLLIGLSAALV-FNTLQH 762
QY 758 --WTCNVISQVTHRYOAAAFORVLDODIELDPEIQISGALTQSLSALPTOLQELISAN 815
Db 763 SFW--DIVGEMLTAKVRKEMLSAVLKNEAMFWQEBENESARIAARLADANNVRSAGDR 820
QY 816 FLIYIVWGQHRLEOCSTTSMWETGPGGVNCTSTPAFGW----- 855
Db 821 ISVIV-----QNTALML-----VACTAGFVLQWRLLVLYAVFPVVAATVILQ 863
QY 856 ---LPQNSRDRSRLKGLCKRCKWACKRSYRDPDRLIFDSRRRPSPTVLGHVYEQGLA 912
Db 864 KHEMTGFGSDLEAAHAHAGTQLAGEAIVNRTV-----AAPNSEAKIVRLVYANLEPPLK 917
QY 913 KTIQSFNFGRC--FGFHLUSQSMFLALALGFCIAVDNML-----QVSTTOLNFISSSWA 964
Db 918 R---CFWKGOIAGSGYGAQFCLYASVALGLWYA--SMLVKHGISDFSKTIRVFM---V 968
QY 965 FCLPQAAAOYLAYSTSTFKARSAANYILWRLTKPTIRETEENKKGP--VGCGPVDLE 1022
Db 969 LMSANGAAETLTUADPFIKGGQAMRSVFELDRKTEIEPDDPTTPVDPDRUG-EVELK 1027
QY 1023 DIEFRYRDRSARVLRGVSMTEPQFVAVVGASGCGKSTLIALSERFYDPTSGRISPAH 1082
```



Db 1028 HIDEPSRDPDIOIFRDLGSLRARAGKTALVGPSCGKSVISLIQRFFTEPSSGRVMIDG 1087  
Qy 1083 ENTAEMSPRLRYGHMSLYOOEPTLYOGSVRENTVTLAEAELEBELCQGRLPARPMLWLS 1142  
Db 1088 KDIRKYNLKAIRKHTAIVPQECFLFGTTIYENIAYGHECATEAEIIQAATLASAHKFT-- 1145  
Qy 1143 SLYQKALKRLAAGQMGQSGGQRIARALIRNPKLILLDLATSDALDQSERIVQAAL 1202  
Db 1146 SALPEGYKTYVGGRGVOLSGGQRIARALVRKAEIMLLDEATSDALDAESRSVQAL 1205  
Qy 1203 DEASTSTTAVAHRLSTIRNVDFVFPANGRIAEFTGTHAEL 1244  
Db 1206 DQACSGRTSIVVAHRLSTIRNAHVIAVIDDGKVAEQGSHSL 1247

## RESULT 5

US-09-540-209B-6953  
; Sequence 6953, Application US/09540209B  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR  
; FILE REFERENCE: 2709,1001-001  
; CURRENT APPLICATION NUMBER: US/09/540,209B  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 6953  
; LENGTH: 612  
; TYPE: PRT  
; ORGANISM: B.fragilis  
US-09-540-209B-6953

Query Match 10.6%; Score 682.5; DB 5; Length 612;  
Best Local Similarity 30.6%; Pred. No. 2.3e-45;  
Matches 189; Conservative 104; Mismatches 253; Indels 71; Gaps 11;  
Qy 48 LNVIALCAIGSGASI-PLMSTIF-----GSGFTKNNKNSGDS 86  
Db 26 LNLISALNIFPSLSLAPILQILFKVDYDFPIPDVTWISPKDILLNNFYVSA--- 82  
Qy 87 PEAPKADVDFVLMFWYLPBIGFVL-----YVSTAAITISAIRP--TRTFLRVFLECT 138  
Db 83 --MIQANGASFTLMILGLFL--VWMTFFKTCYFASSAVMI-PLRTGVVVRDIRVMVYDKV 137  
Qy 139 LRQEVWHRFQSGNGAIATVTTNGNRIGTGAERKLVTVQALSMFTSAFVVALASQWKL 198  
Db 138 LRLPLGPFSEERAGDIARMGSDVGEIENSVTSSIDMLVKNPILIIAYETTLIVTSWOLT 197  
Qy 199 LITMSVIPALFVLTGICIAIDAAQEARITRIYSRAAVALAEVLSSRTVHAFYAKMVE 258  
Db 198 LFTLLVPLMGWVGTVGKLRQSLAQAKNSDTMSQLEETLGLRIIKATIAENKIN 257  
Qy 259 KYDVFLOAQHQEGKSPNGVLFSTEFYCIYAALAPKPFMYQNGEVADVGVTV 318  
Db 258 RFLKCSNEFRDATNKVATQALAHFPMSEFLGTTMIVILWPGGSLILSSHSDASTPMV 317  
Qy 319 VLSVTLAATSIEMLAPSGSVYQRRIFGSELFSDIDRPTQLDPLD-----PSG 366  
Db 318 YLTILY-----SVINPLKFEKAGYNIPKGLASMERVDKILKAENKIVEINP 365  
Qy 367 KQPEGCLGQIEIONLAFAYPSRPSAQVLRDNFLTPAGKRTALVAGSGSKSTMVGLLER 426  
Db 366 KPLNGLEQVEFRDIDSFSDGK--KEVLQHNLNTPVKRTIALVQSGSKSTLVLLPLR 423  
Qy 427 WLPSSGRILLDGLGQNVKWLRSRLRVQEPVLRGTIFQNTANGFMDEQRDLPRE 486  
Db 424 YHDVQEGTIIIDGVNKKDVRISDLRSLGNVYNQEAILEPNTFFNIAFGV-----EN 475  
Qy 487 QMELVQKACKASGVDFINELPNGYETTEVGERAGALSGGQRQRIATNSIDPKILL 546  
Db 476 ATMEQVTEAAKIANAHDFIMEXEDGYHTNIGDRGSKLSGGQRQRISTARILKNPPILT 535

Qy 547 DEATSALDPKAEKVQAEALNRYSKORTTLVIAHKLATVKSAGNIANVISOCKIVEQTHHE 606  
Db 536 DEATSALDTESELRVEALERLMKTRTTIAJHRLSTIKNADEICVLYGEIVEGRCKHEE 595  
Qy 607 LIEFGCHYAALVRAQDL 623  
Db 596 LIAKNGYKRLNDMQSL 612

## RESULT 6

US-09-573-655B-1015  
; Sequence 1015, Application US/09573655B  
; GENERAL INFORMATION:  
; APPLICANT: SOLOVIEV, Victor and TROUKHAN, Maxim  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 2750-0876P  
; CURRENT APPLICATION NUMBER: US/09/573,655B  
; CURRENT FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 3281  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1015  
; LENGTH: 578  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-573-655B-1015

Query Match 8.5%; Score 546.5; DB 5; Length 678;  
Best Local Similarity 28.2%; Pred. No. 9.5e-35;  
Matches 181; Conservative 122; Mismatches 273; Indels 65; Gaps 20;

Qy 10 PNQGSSESHPPPEKETEGSIDYLRIFRYADKYDNTLVNI-ALICAGSGASLPLMSI 68  
Db 70 PNQDOTKAS----SKILKRTISSVLMW---KDNPELRFVRVIAALACLIGAKFLNQVVF 122  
Qy 69 IFGSFTKFNKNNYAG---DGSP---EAFKADVDFVLMFWYLPBIGFVLTVYVSTAAITIS 122  
Db 123 LFKLSIDLLSSYSSTITDSNPVLLAAF-ATPSSVLIGYIARSGSSAFNELRTAVFSKV 181  
Qy 123 AIRTRTLRRVFL---ECTLRQEVWHRFQSGNGAIATVTTNGNRIGTGAERKLVTVQ 178  
Db 182 SLRTIRSVKRVLSHLHDLDR---YHLNRET-GAL-NRIIDRGSRINILSAMFNVV 236  
Qy 179 ALSMFFSAFVVALASQWKLALITMSVIPALFVLTGICIAIDAAQEARITRIYSR----- 232  
Db 237 PTILEISNVVTGILAYNFG-----PVFALITSLSYGSYIAFTLVVTVQYRTKFRKMN 287  
Qy 233 -----AAVLAEEVLSSRTVHAFYAKMVEKYDVFLOAQHQEGKSKSPNGVLFSTEF 287  
Db 288 QADNDASTRAIDSLNVYTVKYVFNNEYEARKYDLDLLGRYEDAAALQTKSLAFDFGQSF 347  
Qy 288 CIYAAIALAFKPFMYQNGEVADVGVTVVLSVTLAATSIEMLAPSGSVYVQRRIFG- 346  
Db 348 IFSTALSTSMVLCSSQGTMGEMT-VGDL--VWVNGLLFQLSLPLFLGG--VYRETVQGL 402  
Qy 347 ---SELFSIDKPTQLDPLDPSGKQPEGCL--GQIEIONLAFAYPSRPSAQVLRDNFLI 401  
Db 403 VDMKSLFQLEERSDIGDKDTETKLPPLVLRGSGISFENVHFSY--LPERKILDGISFEV 460  
Qy 402 PAGTITLVAGASGSGKSTMVGLERWLPSSGRILLDGLGQNVKWLRSRLRVQEP 461  
Db 461 PAGKSAVATVSGSGSGKSTILRMIFRFDTSNGNRIDGQDIKEVLTESLSCIGVYPOST 520  
Qy 462 VLFRTTIFONTANGFMDEQRDLPREQMELVQKACKASGVDFINELPNGYETTEVGERAG 521  
Db 521 VLENFTIFNHIYGNLSATEE-----EVDYDAARRAVIHDT-IMKFPDKYSTAVGERGL 572  
Qy 522 ALSGGQRQRIATNSIDPKILLDDEATSDALDPKAEKVQAEALNRYSKORTTLVIAHKL 581  
Db 573 MLUSGGEKORVALAKAFKLSPAILLCDENATLSDSKTEAEINKTFRSLASNRTCIFIAHRL 632



```
; Sequence 726, Application US/10102806
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PLC1
; CURRENT APPLICATION NUMBER: US/10102.806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/945,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05981
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 726
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (108)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (383)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (688)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (690)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; US-10-102-806-726

Query Match 7.9%; Score 512; DB 6; Length 690;
Best Local Similarity 28.8%; Pred. No. 4.6e-32;
Matches 151; Conservative 99; Mismatches 219; Indels 56; Gaps 11:

Qy 61 ASPLMSIIFGFTKFNNGSGSPFAKADVDHVFVLFYFGKFKVLTIVTAAAT 120
Db 192 SSLGEMAIFF--FTGLDWTLDQGSADTFTRNLTLMSI---LTIASAVLEFVGDIYN 245
Qy 121 ISAIRTRTLRRVFLFECTLRQEVNHFQKSNAGAIATXVTNGNRIOGTGAERLVFTVQAL 180
Db 246 NTMGHVHSHLQGEVFGAVLRQETEPFOQNTGNIMSRVTEDTLSLSLENLSLFLWYL 305
Qy 181 SMFFSAFVVALASQMKLALITMSVIPALFVLTGICIAIDAAQEARITRIYSRAAVLAEEV 240
Db 306 VRGLCLLGLMNGSVSLTWVTLITLPLFLPKVKGWYQLLEVQVRESLAKSSQVAIEA 365
Qy 241 LSSIRVHAFYAQKKWVEKYDFVLOQAHOEGKKSPNNGVLFSTEYFCIYAAALAFWK- 299
Db 366 LSAMPTVRSEAFEEGAKFKREKKEIKTLNKEA-----VATAVNSWTT 410
Qy 300 -----GFRMYQNGEADVGKFTVTVLSVTLAATSI--SMLAPS--GSVYVQRRIFGS-----ELFS 351
Db 411 SISGMLLKVCILYIGGLVTSQVAVSS--GNLVTFVLYQMOTQAVFVLLS-----IYPRVQK 465
Qy 344 IFGS--ELFSIDKPTQLDPLDPG--KOPEGCLGQIEIONLAFAPSPRPSAOLRDNFLT 400
Db 466 AVGSSEKIFEYLDTRP---CPPSGLLTPLHLEGLVQFQVDSFAPNRPDVLVGLGTF 522
Qy 401 IPAGKTTALVAGSGKSTWVGLLERWYLPSSGRILLDGLGQYVNVKWLRSRIRLVQOE 460
Db 523 LRPEVTALVPGNGSKSTVAALLQNLQYPTGGQLLDGKPLPQYEHRYLHRQVAVGOE 582
Qy 461 PVLFRTIFQNTANGFMDEQRDLPRKQELVQKACKASNDGVFINELPNGYETEYGERA 520
Db 583 POFVGRSLQENIAYG-----LTQKPTMEBITAAAVKSGAHSFISGLPQGYDTEVDEAG 635

; Sequence 726, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6291
; LENGTH: 808
; TYPE: PRT
; ORGANISM: B. fragilis
; US-09-540-209B-6291

Query Match 7.9%; Score 509; DB 5; Length 608;
Best Local Similarity 27.5%; Pred. No. 6.5e-32;
Matches 170; Conservative 109; Mismatches 256; Indels 84; Gaps 16;

Qy 25 KETEGSISDYLRIFRYADKYDWTLVNIALICAGSGASLPLMSIIFGFTKFNNGNSGD 84
Db 9 KKREG-----LSRLEFETAGOKKSLILLAGLSA---GSACMLIPYAIYRILVELLNH-- 59
Qy 85 GSPEAKADVDHVFVLYF--TGKFLVTVSTAAITISAITRTRLRVFELECTLRQ 141
Db 60 -SRELSIDETNNIRWGIAFGGLIGLLLYASLSMSHVAAYRILYCLIRTEHIGRL 118
Qy 142 EVNHFQKSNAGAIATXVTNGNRIOGTGAERLVFTVQALSMEFFSAFVVALASQMKLALIT 201
Db 119 PLCYNGTSTGAIKKTEQNEKIEFNIAITIPDLVNVNATVNVNLFIFSLDGLWAGVC 178
Qy 202 MSVIP-AIELVTCICIAIDAAQEARITRIYSRAAVLAEEVLS-----242
Db 179 LAVIVSIFLOPSNFMKKARE---FTRIYN---AOEQMSASAVQYVRGMPVVKFGQ 231
Qy 243 SIRTVHAFYAQKKWVEKYDFVLOQAHOEGKKSPNNGVLFSTEYFCIY--AAIALAPWK 300
Db 232 SVRSFRQFNAIEYAKTYALKVCDTYESG-----MTYFTVLLNSIVTFLPVG 279
Qy 301 FRMYQNGEADVGKFTVTVLSVTLAATSI--SMLAPS--GSVYVQRRIFGS-----ELFS 351
Db 280 ILLMONDS-----RSLTLAAVWLFILPGVAVSPYKLVLSGSTREINEGVS 328
Qy 352 IIDKPTQLDPLDPGSKOPEGCLGQIEIONLAFAPSPRPSA--OVLRFNLTIPAGKTTA 408
Db 329 RIDRILENQVSEFACPKIPATYDIEFRHVSFSEYENKEQATETALHDLCTAPQKITA 388
Qy 409 LVGASGSGKSTWVGLLERWYLPSSGRILLDGLGQYVNVKWLRSRIRLVQOEVLFGCTI 468
Db 389 FVPGSGKSTVANKLIPREFWDEQGEILIGNVNVKDIATEQMDLVISFVQDTFLFDL 448
Qy 469 FQNTANGFMDEQRDLPRKQELVQKACKASNDGVFINELPNGYETEYGERAGALSOGOR 528
Db 449 YENIAGVSSKATRD-----TVIAAARAAQCHIEFKLPNGYETRIGDKGVPLSGGEA 500
Qy 529 QRTAIRSISDPKILLDDEATSDLPKAEKVQVQALNRVSKDRFTTVIAHKLATYKASG 588
Db 501 QRVCVARAILKNAPILVDEATAFADPENYEYKMQQALASILKDKTVIIIAHRLSSIVSSD 560
Qy 589 NIAVISQKGTVEQSTHHEL 607
Db 561 RIIVLKDGRAVQCGRHEEL 579
```

```
RESULT 11
US-09-540-209B-8625
; Sequence 8625, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8625
; LENGTH: 740
; TYPE: PRT
; ORGANISM: B. fragilis
US-09-540-209B-8625

Query Match      7.7%; Score 498; DB 5; Length 740;
Best Local Similarity 27.3%; Pred. No. 6.2e-31;
Matches 183; Conservative 111; Mismatches 263; Indels 114; Gaps 21;

QY 5 SEKPRPDGSESSSHPPPEKETEG---SISDYLRIPI-RYADKYDWTNLNVALICAGS 59
Db 134 STKNGEKGAALEPCPGQGGDEESGRSRFFLKLSPY--RKQLIQLILGM-- 189
QY 60 GASLPLMSIIFGSFTNKNFNYSNGDGSPEAFKADVDHVFVLFVYL 104
Db 130 -LTASLLQLIPFLTOSLDVIGIRDN-----LNFITLILISQLVISVLSV 236
QY 105 -FIQKFLVTVYSTAAITISAIRTRTLRRVFELECTLRQVWHFQKSNALATVVTNGN 163
Db 237 EFIRSMWLMHNT-RINISLISD-----FLAKMLKPLHYDFDKMIGDMQIRG 287
QY 164 RIOTGIAEKVETV-QALSFFFAFVVALASOKKLALITMSVIPAFIVTIGICATA 222
Db 298 RIESFLGSSISILFSVNFPIAPVLA--YNNGLIAFLVGNLSLILCMILVFMYRR 344
QY 223 EARTIRYSRAVLAERVLSSITRTVHAFAQKKMVEKYDVFLQQAHOEGKKSPNNGVLF 282
Db 345 ELDIRRE-----QAAGSQSSLIOLVTA-----OEIKLNCKEOKR----- 381
QY 283 STEYFCIYAAIALAFMKGFMYONGEVADVGVKF-----TVVLSVTLAATSISMLAPSGS 337
Db 382 -WONERIQVKLFKISVKGLAL---GOVQOVGSVFENQTTNIVISIAAKSVVEGNMTLGM 437
QY 338 VYQORRIFG---SELFSDIKPTQLDPLDPGCKOPEGCLGO----- 375
Db 438 MMSLTYYIGOLSGPIGSFIFGFAOQLQDAKISLERLNEIHGQKDERQDITASKLTVLPERD 497
QY 376 IEIONLAFAYPRPSAQVLRDFNLTPAGKTTALVCASGSGKSTWGLLERWYLPSSGRI 435
Db 498 IRIENLSFSYDGDADRDYVLNDVNLNTPHEKHVTAIVGSGSGKTTLIKILMLGFTYTPNKGDI 557
QY 436 LLDGLEGOYNVWKLRSRIRLVQOEPVLPFRGTIFQNIANGFMDQORDLPREKOMELVQKA 495
Db 558 KIGETPLDVNVNPHLWRAKSGSVNQDGFILSDTIANNIAGV--EEQVDVVER-----PRHA 609
QY 496 CKASNGDVFTNELPNGVETVGERAGALSGGORORTAIARSIISDPKILLDEATSAIDP 555
Db 610 VTVANIRDFIDSLPLGYNTKIGNEGNCISQGORQALLARAVYKNPEFLPFDDEATNALDA 669
QY 556 KAEKVQOALNRVSKORTTIVIAHKLATVKSAGNTAVISQGIKVEQGTTHHELIEFGCHYA 615
Db 670 NNEIREIMEHLHTYRKTVVVAHRLSTVRDADKIIVLDRGAVABEGTHRELTEKKGLYY 729
QY 616 ALVRAQ-DLGA 625
Db 730 QLVNRNOLELGS 740
```

RESULT 12
US-09-540-209B-5749

```
; Sequence 5749, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDE
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 5749
; LENGTH: 735
; TYPE: PRT
; ORGANISM: B. fragilis
US-09-540-209B-5749

Query Match      7.6%; Score 490; DB 5; Length 735;
Best Local Similarity 27.6%; Pred. No. 2.6e-30;
Matches 183; Conservative 117; Mismatches 253; Indels 110; Gaps 26;

QY 3 DESE-----KPRPDGSESSSHPPPEKETEGSISDYLRIPIRYADKYDWTNLNVALICAG 58
Db 136 DEQQVLLAVEPTTDFKNRAE---QEQKRNFSLSILKTFPPYKKSFGILFIMLVTVVL 192
QY 59 GASLPLMS---IIFGSFTNKNFNYSNGDGSPEAFKADVDHVFVLFVLFVLFVLFV 115
Db 193 QG-KLPEFSKAVIDVGKTSDRNFNN-----VLIGNICILLSVNIFNVLRDWTLLHI- 244
QY 116 TAAITISAIRTRTLRRVFELECTLRQVWHFQKSNALATVVTNGNRIQGTGIAEKLVF 175
Db 245 TARNVIALISD-----YLKLMKLPVTFENKLLGDLQRAODH-ERIRSFINN--- 292
QY 176 TVALSMPFSAFVVALASOKKLALITMSVIPAFIL---VTGIC-----LAIDAAQEARIT 227
Db 293 --NSLALIFSLTFAVS--LILLYNTIIFYIFLFGSVLYACVLLFLSIRKLDWEYF 348
QY 228 RIYSRAVLAERVLSSITRTVHAFAQKKMVEKYDVFLQQAHOEGKK-----KSPNNGVLF 282
Db 349 ELLSKNGSYVETVSTIQDIKIYNDKYRRKWEETQARLYHVNRKRVLAITNAQNLGAQF 408
QY 283 -----STEYFCIYAAIALAFMKGFMYONGEVADVGVFTVVLVSVTLAATSISMLAPS 335
Db 409 IENTKNMAIVFFFCAMAVI-----KGEI-----TFGIMISTQFI---IGML--N 446
QY 336 GSVV-YORRIFGSELFISI---IDKPTQLDPLD-----PSGKQPEGCLGOIEIQ 379
Db 447 GPLVQFINFVVSQAQYAKISFLRINEIROLENEDELLSIGSTILPERK-----TILE 499
QY 380 NLATAY-PSRPSAQVLRDFNLTPAGKTTALVCASGSGKSTWGLLERWYLPSSGRILLD 438
Db 500 NIHFQYTPNSP--LVLNRYLQIPENKITAIVGSGSGKSTLLKLLVRLYRPSHGEIKMD 557
QY 439 GLELGOYNVWKLRSRIRLVQOEPVLPFRGTIFQNIANGFMDQORDLPREKOMELVQAKCA 498
Db 558 KMWVSAINLRQWRNMCMQVMQDKIFSDTILNNIV--LDEQINVTYR-----LREVCR 609
QY 499 SNGDVFTNELPNGVETVGERAGALSGGORORTAIARSIISDPKILLDEATSAIDP 558
Db 610 AQIEDEINAMPKGETTIGETGRGLSGGQKQRLIARALYROPKFLPMDXATNSLDSINE 669
QY 559 KVVQOALNRVSKORTTIVIAHKLATVKSAGNTAVISQGIKVEQGTTHHELIEFGCHYA 618
Db 670 RKIYNALNNAFEQRTVVVVAHRLSTIRNADQIVVLQKGFIVETGTHEILMEKKKHGFELV 729
QY 619 RAQ 621
Db 730 SSQ 732

RESULT 13
US-09-540-209B-6190
; Sequence 6190, Application US/09540209B
; GENERAL INFORMATION:
```

```
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6190
; LENGTH: 574
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-6190

Query Match      7.3%; Score 468; DB 5; Length 574;
Best Local Similarity 27.7%; Pred. No. 9e-29;
Matches 153; Conservative 110; Mismatches 220; Indels 70; Gaps 17;

QY 110 VLVYSTAAITSAIRTRTTRRRVLE- - - - -CTLR- - - - -QEVWHF--DKOSNG 152
DB 51 VITAVGKMLISIVGLIFSANYVSSHASIGFGDRLTGFLGKIQLQSFEDIRESTA 110
QY 153 AJATXVTNCNRIOTGIAEKLVFTVOALSMFSAFVVALASOMKALITWSVIP- - - - -AI 208
DB 111 SLITRIFSDISRIQOIVIMSRMLRSPLMLVMAVEFVVRINLELAGVLLAIPILGFSV 170
QY 209 FLV--TGICIAIDAAQERITRIYSRAAVIAEVLSSITRVHAFYAQKKWVEKYDVPELOQ 266
DB 171 FFILRKGFFFLKVOQKV- - - - -DQLEVVRENINIRVKSFVREDFEAHRF- - - - - 218
QY 267 AHQEGKKKSPNGV- - - - -LFTSEYFCIIYAATAALAFWKGFWMYONGEVAADVGVKFT 317
DB 219 - - -KDKESLRVTIVHASNITVSFPVMQVLMNLSIIAILMGMGHKVMYTGEL-KVGLIS 274
QY 318 VWL- - -SVTLAATISMLAPSGSVVYQRRIFGSE-LFSIID-KPTQLDPLDPSKQP--E 370
DB 275 FVNYLGQVLSMLMSLIMS- - - - -YARASASKRIEVLDTQPSLTD--TPEGMQSTRE 328
QY 371 GCGQTEIONLAFAYSPRSQAQLVRLDNLNIPAGKTTALVGASGSKSTWYGLERWYLP 430
DB 329 IEKGEIAFEKVSFRY-GGGETVLRNLSFHIRPGETVATAGATGSAKSLVQLIPRLYDV 387
QY 431 SSGRIILGLELGOYNVAVLSRIRLVQOEPVLFRTIFONFANGFMDQERDLPREKOME 490
DB 388 SAGEIRIDGIPQVDYNLRELHARICGMVLQKNELFTCTIAENLRWCKPDATQ- - - - -E 439
QY 491 LVOKACKASNGDVFNELPNGYETEVEGERAGALSGGQRORIAIARSIIDPKILLDEAT 550
DB 440 ELEVAARAAEAHEFICSLPAGYDTLLGRGGINLSGGOKORICIAALLRKKPILLDOST 499
QY 551 SALDPKAEKVQOEAALNRVSKDRTWLVIAHKLATVKSAGNIAVISOKKIVEOGTHHELIEF 610
DB 500 SAVDSFTELRIRNNLNAMLRTVTLITQRIYTMQSANRVILLDDGEIESIGTPEELLER 559
QY 611 GCHVAALVRAQDL 623
DB 560 SEMYEIYYSQKI 572

RESULT 14
US-09-540-209B-8427
; Sequence 8427, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8427
; LENGTH: 740
; TYPE: PRT
```

```
; ORGANISM: B.fragilis
US-09-540-209B-8427

Query Match      7.1%; Score 457; DB 5; Length 740;
Best Local Similarity 26.4%; Pred. No. 9.3e-28;
Matches 170; Conservative 115; Mismatches 251; Indels 108; Gaps 23;

QY 20 HPPPEKETEGSISDYLRIFRYADKYDWTNLNITALICATGSGASPLPMSIIFGFTNKNFN 79
DB 159 HNMNEMETRASFSHFMYLAPHRYLGQV-IVGMVAGILIGLSPFTS- - - - -OSIVD 210
QY 80 YNSGDGSPFAKADVDHFVLFVY- - - - -YLFICKFVLYTVSTAAITTSAIRTRTLRRV- - - - - 133
DB 211 FGIQSGN- - - - -IQFVNTMLIAGMILAFSSMASDFIQSRLMLYVYSRINMGV 258
QY 134 - - -FLECLTRQBVHFDKQSGAIGATXVTYTNHRIOTGIAEK- - - - -LVFTVQALSMF 183
DB 259 SDFLRKTKMSLPTTFEERKMYVDLLTRIDDHG-RIQSFIMSTFLGIFINILFLVYSLML 317
QY 184 F- - -SAFVVALASOMKALITMSVIPAFELVTGICIAIDAAQERITRIYSRAAVLAEEV 240
DB 318 YVESNMFLVPMIGN- - - - -TVYTGWIFELFKQKKLDNOLFGCRATNNDL 363
QY 241 LSSIRTHAF- - - - -YAOCKM- - - - -VEKYDVF- - - - -LQAHQEGKKKSPNGVLFST 284
DB 364 LLELVNVEIKINNIANKRKKWELSRFKIYGLRVKNNMNDLQIATGASFIHLOGLFIT 423
QY 285 EYFCIYAAIALAFWKGFQM- - -YQCEV-ADVGVFTVVLVTLAATISMLAPSGSVY 340
DB 424 -YIAALNVICTWTLGMMMAAQILGQNAPIKSMIGVHSLQPARISLKV- - -NEVIM 479
QY 341 QR- - -RIFGSELSFIIDKPTQLDPLDPSGKQPGSCGQIEIONLAFAYPSPSA-QVLRDF 397
DB 480 EEPESISESKSYIEIK- - - - -GVKKDDLFY- - - - -NPNLUNKVLDNI 518
QY 398 NETIPAGKTALVGASGSKSTWYGLERWYLPSSGRILLDGLGOYNVYKWLRSIRLV 457
DB 519 NLEIPEGKITAIVGESGKTLKLLRKYPTNGEIEVGVPLDNLIDYRNNSCGAV 578
QY 458 QOEPVLFRTGTFQNIANGFMDQERDLPREKOMELYOKACKASNGDVFNELPNGYETEVG 517
DB 579 LQDGLKFLNDTLYNIT- - -LEDEEMNV- - -NOKQLV-KAIQLANAENFINARPLKLYPLG 632
QY 518 ERAGALSGGQRORIAIARSIIDPKILLDEATSAIDPKAEKVQOEAALNRVSKDRTLVI 577
DB 633 TNGSGLSQGOKORILIAARLYKNPDEFILDEATNSLDTNNEKQISKNLETILEGKTAIVI 692
QY 578 AHKLATVKSAGNIAVISOKKIVEOGTHHELIEFGCHYAAVRAQ 621
DB 693 AHRSLTVRNAHNVMEKGIQVGGTHQELINLKGIIYDYLSSQ 736

RESULT 15
US-10-076-157-4
; Sequence 4, Application US/10076157
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Suelberger, Harald
; APPLICANT: Josefken, Hans Wolfgang
; APPLICANT: Doval, Jose Luis Revuelta
; APPLICANT: Jimenez, Alberto;
; APPLICANT: Garcia, Maria Angeles Santos
; TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii and the us
; TITLE OF INVENTION: in
; TITLE OF INVENTION: in
; FILE REFERENCE: 48684DIV
; CURRENT APPLICATION NUMBER: US/10/076,157
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 09/212,247
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Wordperfect v. 6.1
```

```

; SEQ ID NO 4
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Asbya gossypii
US-10-076-157-4

```

```

Query Match      6.7%; Score 434; DB 6; Length 475;
Rest Local Similarity 28.6%; Pred. No. 2.9e-26;
Matches 131; Conservative 79; Mismatches 176; Indels 72; Gaps 13;

QY 197 LALYMSVIPAFILVGTGICAIID--AAQBARIT-----RIYS----- 231
Db 12 LSAMVFHIIPIITEISWGCILITVQFGASPAATFTMLLYSIFTFRTAMRTFRERDAN 71
QY 232 -----RAAVLAEBVLSIRTVHAPYAAQKMWKEVDVFLQQAHOEGKKKSPNNGVLSTFYF 287
Db 72 KADNKAASVALDSLINFEAATVFNENKYLADKYHTSLMKYRDSQTKYSQSLAFLNTGQNL 131
QY 286 CYAAIALAFWKGRFYONGEVADGVKFTVVLVSVTLAATISMLAPSGSV--VYQRI 344
Db 132 IFTALTAMMYMACNGVMQGSULT-VGDLVLINQLVQLSPVPLNF--GSIYRDLKQSL 187
QY 345 FGSFLF-----SIIDKPTQID-----PLDPGKGQEGCIGQIBQNLAFAYPSRPSAQ 392
Db 188 DMESLFLQKNQYTIKNSPNAQNLPIHKPLD-----IRFENVITFGYD--PERR 233
QY 393 VLRDNFLTIPAGTKTALVAGSGCKSTMVGLLERYWLPSSGRILLDGLGELQGVNWKLR 452
Db 234 ILNNVSTTIPAGMKTAIVGSPGCKSTILAVFRPEYEPQGRILVGGTDIROLDLLSLRK 293
QY 453 RIRLVQOEPVLFRTGTFIORIANGFMDQRDLPREKOMELVQKACKASNGDVFINELPGY 512
Db 294 AIGVVPQOTPLFNDTIWENKRGNISDD-----EILRAIEKAQUTK-----LLQNLPKGA 345
QY 513 ETEYGERAGALSGGQRQRIATARSITSDPKILLIDDEATSLDPAKAEKVQOEAL--NRVSK 570
Db 346 STYVGERGLMSGGERQRIATARVLLKDALPFDDEATSLDTHTEQALLHTIQOEFSSN 405
QY 571 DRTTLVLAHLKATVCKSAGNIAVISQKTVPGQTHHELI 608
Db 406 SKTSVYVAHLRTIADADKIIVLEGVSVRGEQTHSLL 443

```

Search completed: April 27, 2002, 07:58:51  
Job time: 752 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 07:55:22 ; Search time 389.72 Seconds  
(without alignments)  
1140.694 Million cell updates/sec

Title: US-09-882-694A-11

Perfect score: 6455

Sequence: 1 MADESEKPRNPQDSESSSH.....LQRLGRYEMCIAQLDQA 1263

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/paa/US081\_COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/paa/US082\_COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/paa/US083\_COMB.pcp.\*  
7: /cgn2\_6/ptodata/2/paa/US084\_COMB.pcp.\*  
8: /cgn2\_6/ptodata/2/paa/US085\_COMB.pcp.\*  
9: /cgn2\_6/ptodata/2/paa/US086\_COMB.pcp.\*  
10: /cgn2\_6/ptodata/2/paa/US087\_COMB.pcp.\*  
11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pcp.\*  
12: /cgn2\_6/ptodata/2/paa/US089\_COMB.pcp.\*  
13: /cgn2\_6/ptodata/2/paa/US090\_COMB.pcp.\*  
14: /cgn2\_6/ptodata/2/paa/US091\_COMB.pcp.\*  
15: /cgn2\_6/ptodata/2/paa/US092\_COMB.pcp.\*  
16: /cgn2\_6/ptodata/2/paa/US093\_COMB.pcp.\*  
17: /cgn2\_6/ptodata/2/paa/US094\_COMB.pcp.\*  
18: /cgn2\_6/ptodata/2/paa/US095\_COMB.pcp.\*  
19: /cgn2\_6/ptodata/2/paa/US096\_COMB.pcp.\*  
20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pcp.\*  
21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pcp.\*  
22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pcp.\*  
23: /cgn2\_6/ptodata/2/paa/US100\_COMB.pcp.\*  
24: /cgn2\_6/ptodata/2/paa/US101\_COMB.pcp.\*  
25: /cgn2\_6/ptodata/2/paa/US102\_COMB.pcp.\*  
26: /cgn2\_6/ptodata/2/paa/US103\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	6453	100.0	1263	17	US-09-351-224-11	Sequence 11, Appl
2	6453	100.0	1263	17	US-09-351-224E-11	Sequence 11, Appl
3	6453	100.0	1263	17	US-09-351-823-11	Sequence 11, Appl
4	6453	100.0	1263	20	US-09-677-488-11	Sequence 11, Appl
5	6453	100.0	1263	20	US-09-677-488A-11	Sequence 11, Appl
6	6453	100.0	1263	20	US-09-677-682-11	Sequence 11, Appl
7	6453	100.0	1263	20	US-09-677-682A-11	Sequence 11, Appl

8	6453	100.0	1263	20	US-09-677-682B-11	Sequence 11, Appl
9	6453	100.0	1263	22	US-09-882-694-11	Sequence 11, Appl
10	6453	100.0	1263	22	US-09-882-694A-11	Sequence 11, Appl
11	2270	35.2	1308	1	PCT-US98-27499-2	Sequence 2, Appl
12	1956	30.3	1349	10	PCT-US86-15734-2	Sequence 2, Appl
13	1940	30.1	1334	7	PCT-US98-27300-2	Sequence 2, Appl
14	1780	27.6	1307	7	US-08-395-246A-2	Sequence 2, Appl
15	1656	25.7	1283	20	US-09-672-810-4	Sequence 2, Appl
16	1654.5	25.6	1280	20	US-09-672-810-2	Sequence 2, Appl
17	1649.5	25.6	1280	20	US-09-584-586-2	Sequence 2, Appl
18	1649.5	25.6	1280	20	US-09-866-866A-2	Sequence 2, Appl
19	1649.5	25.6	1280	24	US-10-072-621-7	Sequence 7, Appl
20	1643.5	25.5	1280	20	US-09-672-810-7	Sequence 7, Appl
21	1643.5	25.5	1280	24	US-10-044-671-2	Sequence 118, App
22	1642	25.4	1241	26	US-60-206-111-118	Sequence 2, Appl
23	1641.5	25.4	1272	1	PCT-US99-20770-2	Sequence 2, Appl
24	1641.5	25.4	1272	16	US-09-208-809-2	Sequence 2, Appl
25	1641.5	25.4	1272	21	US-09-769-097-2	Sequence 2, Appl
26	1641.5	25.4	1272	24	US-10-016-333-10	Sequence 10, Appl
27	1640.5	25.4	1272	1	PCT-US99-20770-4	Sequence 4, Appl
28	1640.5	25.4	1272	15	US-09-156-800-2	Sequence 2, Appl
29	1640.5	25.4	1272	16	US-09-208-809-4	Sequence 4, Appl
30	1640.5	25.4	1272	21	US-09-769-097-4	Sequence 4, Appl
31	1640.5	25.4	1278	22	US-09-805-020-66	Sequence 66, Appl
32	1640.5	25.4	1279	19	US-09-538-082-976	Sequence 976, App
33	1640.5	25.4	1280	19	US-09-584-586-4	Sequence 4, Appl
34	1640.5	25.4	1280	20	US-09-672-810-5	Sequence 5, Appl
35	1640.5	25.4	1280	21	US-09-767-594-2	Sequence 3, Appl
36	1640.5	25.4	1280	22	US-09-817-762-3	Sequence 2, Appl
37	1640.5	25.4	1280	22	US-09-866-866A-4	Sequence 6, Appl
38	1640	25.4	1279	20	US-09-672-810-6	Sequence 2, Appl
39	1631	25.3	1276	11	US-08-784-649-2	Sequence 2, Appl
40	1630.5	25.3	1276	19	US-09-584-586-6	Sequence 6, Appl
41	1630.5	25.3	1276	22	US-09-866-866A-6	Sequence 6, Appl
42	1629.5	25.2	1276	19	US-09-584-586-8	Sequence 8, Appl
43	1629.5	25.2	1276	22	US-09-866-866A-8	Sequence 8, Appl
44	1629.5	25.2	1280	17	US-09-316-167-2	Sequence 2, Appl
45	1629.5	25.2	1280	17	US-09-397-233-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-351-224-11  
; Sequence 11, Application US/09351224  
; GENERAL INFORMATION:  
; APPLICANT: Duivick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351,224  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1263  
; TYPE: PRT  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: p-glycoprotein  
US-09-351-224-11

Query Match 100.0%; Score 6453; DB 17; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADESEKPRNPQDSESSSHPPPEKTEGSDYLRIFRYADKYDWTNLNIALICATGSG 60

Db 1 MADESEKPRNPQDSESSSHPPPEKTEGSDYLRIFRYADKYDWTNLNIALICATGSG 60

QY 61 ASPLMSITFGSFTNKNPNYNSGDSPEAFKADVDFVLFVFLFGKFLVLTYSSTAIT 120  
+++++

Db 61 ASPLMSIIFGSPFNKNYNSGDSPEAFKADVDHFLVFWLFIKGFVLTYYVSTAAIT 120  
Qy 121 ISAIKTRTLRRVFLBCTLRQVWHFDKQSGAIAATVTTNGNRITQGTAEKLVFTVQAL 180  
Db 121 ISAIKTRTLRRVFLBCTLRQVWHFDKQSGAIAATVTTNGNRITQGTAEKLVFTVQAL 180  
Qy 181 SMFSAFVVALASOKKALITMSVIPALFVLTGICIAIDAAQAEARITRIYSRAAVLAEV 240  
Db 181 SMFSAFVVALASOKKALITMSVIPALFVLTGICIAIDAAQAEARITRIYSRAAVLAEV 240  
Qy 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNGVLFSTEFYFCIYAAIALAFWK 300  
Db 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNGVLFSTEFYFCIYAAIALAFWK 300  
Qy 301 FRMYQNGEVADVGKVTYVSVTLAATSISMLAPSGSVYQRRIFGSELSFIIDKPTQD 360  
Db 301 FRMYQNGEVADVGKVTYVSVTLAATSISMLAPSGSVYQRRIFGSELSFIIDKPTQD 360  
Qy 361 PLDPGSKQPEGCLGQIEIONLAFAYPSRPSAQVLRDNFLNIPAGKTTALVGASGSGKSTM 420  
Db 361 PLDPGSKQPEGCLGQIEIONLAFAYPSRPSAQVLRDNFLNIPAGKTTALVGASGSGKSTM 420  
Qy 421 VGLLERWYLPSSGRILLDGLGQYNVWLSRIRLVQOEPVLFRTIFONTANGFMDQ 480  
Db 421 VGLLERWYLPSSGRILLDGLGQYNVWLSRIRLVQOEPVLFRTIFONTANGFMDQ 480  
Qy 481 RDLPREKOMELVOKACKASNGDVFINELPNGYETEVEGERAGALSGGORQRIATARSISD 540  
Db 481 RDLPREKOMELVOKACKASNGDVFINELPNGYETEVEGERAGALSGGORQRIATARSISD 540  
Qy 541 PKILLDEATSAIDPAKRVQOALNRVSKORTTLVTAHKLAVKVSAGNIATVISOQKIVE 600  
Db 541 PKILLDEATSAIDPAKRVQOALNRVSKORTTLVTAHKLAVKVSAGNIATVISOQKIVE 600  
Qy 601 QGTHHELIEFGCHYAALVRAQDLGADQEOHEKTLHEKAAAREAAAGERPALERTHTTATSQ 660  
Db 601 QGTHHELIEFGCHYAALVRAQDLGADQEOHEKTLHEKAAAREAAAGERPALERTHTTATSQ 660  
Qy 661 AGDLERKVPVGTGYSLLKLCILIMFYEQKNLYWCFLLSITVILICAATPPGQALLFSRL 720  
Db 661 AGDLERKVPVGTGYSLLKLCILIMFYEQKNLYWCFLLSITVILICAATPPGQALLFSRL 720  
Qy 721 LTVFELSGHAAQERADFYILMEFFVALGNLVGYFTIGWTCNVISQVVTTRYQAAFPORVL 780  
Db 721 LTVFELSGHAAQERADFYILMEFFVALGNLVGYFTIGWTCNVISQVVTTRYQAAFPORVL 780  
Qy 781 DODIELDIDPEQISGALTQSALPTQLOELISANFLIYIVGQHRLEQCCSTTSLWMEYG 840  
Db 781 DODIELDIDPEQISGALTQSALPTQLOELISANFLIYIVGQHRLEQCCSTTSLWMEYG 840  
Qy 841 PGGVWCTSTPATGWLPONSSRDEARSRLKLCRCWKACRSYRDPORLIFDSRRPCS 900  
Db 841 PGGVWCTSTPATGWLPONSSRDEARSRLKLCRCWKACRSYRDPORLIFDSRRPCS 900  
Qy 901 PTVLGHVQOGLAKIIOQSFWFGRCFGFHLISQSMEFALATGFCIAVONWLVQVSTQUNFIS 960  
Db 901 PTVLGHVQOGLAKIIOQSFWFGRCFGFHLISQSMEFALATGFCIAVONWLVQVSTQUNFIS 960  
Qy 961 SSWAFCLPVQAAAQYLAYSSTFKARSAANYILMLRTLKPTIRETEENKKGKPGVGCPCVD 1020  
Db 961 SSWAFCLPVQAAAQYLAYSSTFKARSAANYILMLRTLKPTIRETEENKKGKPGVGCPCVD 1020  
Qy 1021 LEDIEFRYRDSARVLRGVSMTEIPGQFVAYVAGSGCGKSTLIALSERDYDPTSGRISF 1080  
Db 1021 LEDIEFRYRDSARVLRGVSMTEIPGQFVAYVAGSGCGKSTLIALSERDYDPTSGRISF 1080  
Qy 1081 AHENIAMSPRLYRGHMSLVQOEPVLYOGSVRENVTLALAELESEELCQGRLPARPLMW 1140  
Db 1081 AHENIAMSPRLYRGHMSLVQOEPVLYOGSVRENVTLALAELESEELCQGRLPARPLMW 1140  
Qy 1141 LSSLYQALRKLAAQCHGOFSGGQORQRIATARALIRNPKLLLDDEATSAIDTOSERLVOA 1200  
Db 1141 LSSLYQALRKLAAQCHGOFSGGQORQRIATARALIRNPKLLLDDEATSAIDTOSERLVOA 1200

Qy 1201 ALDEASTRTTIAVAHRLSTIRNVDFVIFPANGRIAETCTHAELOLRGRYVEMCLAQSL 1260  
Db 1201 ALDEASTRTTIAVAHRLSTIRNVDFVIFPANGRIAETCTHAELOLRGRYVEMCLAQSL 1260  
Qy 1261 DOA 1263  
Db 1261 DOA 1263  
RESULT 2  
US-09-351-224E-11  
; Sequence 11, Application US/09351224E  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; TITLE OF INVENTION: Detoxification  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351,224E  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1263  
; TYPE: PRT  
; ORGANISM: Exophiala spinifera  
; NAME/FEATURE: VARIANT  
; LOCATION: 157  
; OTHER INFORMATION: Xaa - Any Amino Acid  
US-09-351-224E-11  
Query Match 100.0%; Score 6453; DB 17; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MADESEKPRNQDSESSHPPEKETEISDYLRIFRYADKYDWTNLVIALICAISG 60  
Db 1 MADESEKPRNQDSESSHPPEKETEISDYLRIFRYADKYDWTNLVIALICAISG 60  
Qy 61 ASPLMSIIFGSPFNKNYNSGDSPEAFKADVDHFLVFWLFIKGFVLTYYVSTAAIT 120  
Db 61 ASPLMSIIFGSPFNKNYNSGDSPEAFKADVDHFLVFWLFIKGFVLTYYVSTAAIT 120  
Qy 121 ISAIKTRTLRRVFLBCTLRQVWHFDKQSGAIAATVTTNGNRITQGTAEKLVFTVQAL 180  
Db 121 ISAIKTRTLRRVFLBCTLRQVWHFDKQSGAIAATVTTNGNRITQGTAEKLVFTVQAL 180  
Qy 181 SMFSAFVVALASOKKALITMSVIPALFVLTGICIAIDAAQAEARITRIYSRAAVLAEV 240  
Db 181 SMFSAFVVALASOKKALITMSVIPALFVLTGICIAIDAAQAEARITRIYSRAAVLAEV 240  
Qy 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNGVLFSTEFYFCIYAAIALAFWK 300  
Db 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNGVLFSTEFYFCIYAAIALAFWK 300  
Qy 301 FRMYQNGEVADVGKVTYVSVTLAATSISMLAPSGSVYQRRIFGSELSFIIDKPTQD 360  
Db 301 FRMYQNGEVADVGKVTYVSVTLAATSISMLAPSGSVYQRRIFGSELSFIIDKPTQD 360  
Qy 361 PLDPGSKQPEGCLGQIEIONLAFAYPSRPSAQVLRDNFLNIPAGKTTALVGASGSGKSTM 420  
Db 361 PLDPGSKQPEGCLGQIEIONLAFAYPSRPSAQVLRDNFLNIPAGKTTALVGASGSGKSTM 420  
Qy 421 VGLLERWYLPSSGRILLDGLGQYNVWLSRIRLVQOEPVLFRTIFONTANGFMDQ 480  
Db 421 VGLLERWYLPSSGRILLDGLGQYNVWLSRIRLVQOEPVLFRTIFONTANGFMDQ 480



QY 481 RDLPREKOMELVOKACKASNGDVFINELPNGYETETVGERAGALSGGORORAIARSISD 540  
DB 481 RDLPREKOMELVOKACKASNGDVFINELPNGYETETVGERAGALSGGORORAIARSISD 540  
QY 541 PKILLDEATSAIDPKAEKVQVEALNRVSKDRITLVIHKLATVKSAGNIAVISOGKIVE 600  
DB 541 PKILLDEATSAIDPKAEKVQVEALNRVSKDRITLVIHKLATVKSAGNIAVISOGKIVE 600  
QY 601 QGTHHELIFGCHYAALVRAODLGADQEQHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
DB 601 QGTHHELIFGCHYAALVRAODLGADQEQHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
QY 661 AGDLKRRKVPVGTGLGYSLLKILIMFYEQKNLYWCFLLSITIVLCAATFPQALLFSRL 720  
DB 661 AGDLKRRKVPVGTGLGYSLLKILIMFYEQKNLYWCFLLSITIVLCAATFPQALLFSRL 720  
QY 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAMFORVL 780  
DB 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAMFORVL 780  
QY 781 DODIELLDIPEQISGALTSQSALPTQLOELISANFLIYVVGQHRLEOCSTSLUMMETG 840  
DB 781 DODIELLDIPEQISGALTSQSALPTQLOELISANFLIYVVGQHRLEOCSTSLUMMETG 840  
QY 841 PGGCVMCTSTPAFCWLPQNSRDEARSKLGLCKRCWACKRSSYRDPDRILFDSRRPCS 900  
DB 841 PGGCVMCTSTPAFCWLPQNSRDEARSKLGLCKRCWACKRSSYRDPDRILFDSRRPCS 900  
QY 901 PTVLGHVEQGLAKIIQSFWRGRCFGPHLSOSMEFLAIALGFCIAVDNKLQVSTTQLNFI 960  
DB 901 PTVLGHVEQGLAKIIQSFWRGRCFGPHLSOSMEFLAIALGFCIAVDNKLQVSTTQLNFI 960  
QY 961 SSWAFCLPVQAAQYLAISFTKARSAANYILWLTKLPTITRETEENKKGPGVGCPCVD 1020  
DB 961 SSWAFCLPVQAAQYLAISFTKARSAANYILWLTKLPTITRETEENKKGPGVGCPCVD 1020  
QY 1021 LEDIEFRYQRDSARVLGVSMTLEPGQFVAYVVGASGCGKSTLIALSERFYDPTSGRISF 1080  
DB 1021 LEDIEFRYQRDSARVLGVSMTLEPGQFVAYVVGASGCGKSTLIALSERFYDPTSGRISF 1080  
QY 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYOGSVRENVTLAEALSEELCOGRUPARPLMI 1140  
DB 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYOGSVRENVTLAEALSEELCOGRUPARPLMI 1140  
QY 1141 LSSLQKALKLAAQRCQMGSGGQORAIARALINPKLILLDEATSAIDTQSERLVOA 1200  
DB 1141 LSSLQKALKLAAQRCQMGSGGQORAIARALINPKLILLDEATSAIDTQSERLVOA 1200  
QY 1201 ALDEASTRTTIAVAHRLSTIRNVDFIVFPANGRIAEETGTHAELQRLGRYEMCLAQSL 1260  
DB 1201 ALDEASTRTTIAVAHRLSTIRNVDFIVFPANGRIAEETGTHAELQRLGRYEMCLAQSL 1260  
QY 1261 DQA 1263  
DB 1261 DQA 1263

RESULT 3  
US-09-351-823-11  
; Sequence 11, Application US/09351823  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fumonisins Detoxification  
; FILE REFERENCE: 5718-56  
; CURRENT APPLICATION NUMBER: US/09/351,823  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1263  
; TYPE: PRT  
; ORGANISM: Exophiala spinifera  
; FEATURE:

i OTHER INFORMATION: p-glycoprotein  
US-09-351-823-11  
Query Match 100.0%; Score 6453; DB 17; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MADESEKPRNQGSSESSHPPEKTEGSDYLRIFRYADKYDWTMLNIALICAISG 60  
DB 1 MADESEKPRNQGSSESSHPPEKTEGSDYLRIFRYADKYDWTMLNIALICAISG 60  
QY 61 ASLPLMSIIFGSFTNKNKNNYNSGDSPEAKADVDHFVLFVFGFVLTIVVSTAAT 120  
DB 61 ASLPLMSIIFGSFTNKNKNNYNSGDSPEAKADVDHFVLFVFGFVLTIVVSTAAT 120  
QY 121 ISAITTTTLRRVFELECTLQEVWHFDKQNGAIAATVTTNGRITOTGIAEKLFTVQAL 180  
DB 121 ISAITTTTLRRVFELECTLQEVWHFDKQNGAIAATVTTNGRITOTGIAEKLFTVQAL 180  
QY 181 SMFFSAFVVALASOMKIALITMSVIPALFVLTGICIAIDAAQEARITRIYSPAALAEV 240  
DB 181 SMFFSAFVVALASOMKIALITMSVIPALFVLTGICIAIDAAQEARITRIYSPAALAEV 240  
QY 241 LSSIRTVHAFYAQKMWKDYDFLQQAHOQKKGKSPNNGVLFSTEVFCYIAAIALAFWK 300  
DB 241 LSSIRTVHAFYAQKMWKDYDFLQQAHOQKKGKSPNNGVLFSTEVFCYIAAIALAFWK 300  
QY 301 FMYONGEADVGVVTVLSTLAATSISMLAPSGSVVQRIKSELSFIIDKPTOLD 360  
DB 301 FMYONGEADVGVVTVLSTLAATSISMLAPSGSVVQRIKSELSFIIDKPTOLD 360  
QY 361 PUDPSGKOPESGICIGIETQNLAFAYPSRPSAQVLRDNLTIPAGKTATLVAGSGSGKSTM 420  
DB 361 PUDPSGKOPESGICIGIETQNLAFAYPSRPSAQVLRDNLTIPAGKTATLVAGSGSGKSTM 420  
QY 421 VGLLEWRVLPSSGRILLDGLGELQYNNVKLRSIRLVQOEPVLFRTIFONTANGFMDEQ 480  
DB 421 VGLLEWRVLPSSGRILLDGLGELQYNNVKLRSIRLVQOEPVLFRTIFONTANGFMDEQ 480  
QY 481 RDLPREKOMELVOKACKASNGDVFINELPNGYETETVGERAGALSGGORORAIARSISD 540  
DB 481 RDLPREKOMELVOKACKASNGDVFINELPNGYETETVGERAGALSGGORORAIARSISD 540  
QY 541 PKILLDEATSAIDPKAEKVQVEALNRVSKDRITLVIHKLATVKSAGNIAVISOGKIVE 600  
DB 541 PKILLDEATSAIDPKAEKVQVEALNRVSKDRITLVIHKLATVKSAGNIAVISOGKIVE 600  
QY 601 QGTHHELIFGCHYAALVRAODLGADQEQHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
DB 601 QGTHHELIFGCHYAALVRAODLGADQEQHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
QY 661 AGDLKRRKVPVGTGLGYSLLKILIMFYEQKNLYWCFLLSITIVLCAATFPQALLFSRL 720  
DB 661 AGDLKRRKVPVGTGLGYSLLKILIMFYEQKNLYWCFLLSITIVLCAATFPQALLFSRL 720  
QY 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAMFORVL 780  
DB 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAMFORVL 780  
QY 781 DODIELLDIPEQISGALTSQSALPTQLOELISANFLIYVVGQHRLEOCSTSLUMMETG 840  
DB 781 DODIELLDIPEQISGALTSQSALPTQLOELISANFLIYVVGQHRLEOCSTSLUMMETG 840  
QY 841 PGGCVMCTSTPAFCWLPQNSRDEARSKLGLCKRCWACKRSSYRDPDRILFDSRRPCS 900  
DB 841 PGGCVMCTSTPAFCWLPQNSRDEARSKLGLCKRCWACKRSSYRDPDRILFDSRRPCS 900  
QY 901 PTVLGHVEQGLAKIIQSFWRGRCFGPHLSOSMEFLAIALGFCIAVDNKLQVSTTQLNFI 960  
DB 901 PTVLGHVEQGLAKIIQSFWRGRCFGPHLSOSMEFLAIALGFCIAVDNKLQVSTTQLNFI 960  
QY 961 SSWAFCLPVQAAQYLAISFTKARSAANYILWLTKLPTITRETEENKKGPGVGCPCVD 1020

```

Db 961 SSWAFCLPVQAAQYLAISTFSTKARSAANYILMLRTLKPTIRETEENKKKGPVGGCPVD 1020
Qy 1021 LEDIEFRYRDRSARVLRGVSMTEIPGQFVAVYGASGCKSTLIALSERFYDPTSGRISF 1080
Db 1021 LEDIEFRYRDRSARVLRGVSMTEIPGQFVAVYGASGCKSTLIALSERFYDPTSGRISF 1080
Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPDLYOGSVRENTVLAELAESELCQGLPAPMLWI 1140
Db 1081 AHENIAEMSPRLYRGHMSLVQOEPDLYOGSVRENTVLAELAESELCQGLPAPMLWI 1140
Qy 1141 LSSLYQAKRLAAQGMQFSGGQORIAIARALIRNPKLLLDDEATSDALDQSERLVQA 1200
Db 1141 LSSLYQAKRLAAQGMQFSGGQORIAIARALIRNPKLLLDDEATSDALDQSERLVQA 1200
Qy 1201 ALDEASTSRTTIAVAHRLSTIRNVDVIFVFANGRIAEFTGTHAELQRLGRYEMCLAQSL 1260
Db 1201 ALDEASTSRTTIAVAHRLSTIRNVDVIFVFANGRIAEFTGTHAELQRLGRYEMCLAQSL 1260
Qy 1261 DQA 1263
Db 1261 DQA 1263

```

# RESULT 4

```

US-09-677-488-11
; Sequence 11, Application US/09677488
; GENERAL INFORMATION:
; APPLICANT: Duwick, Jon
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification
; FILE REFERENCE: 35718/204100 (5718-111A.)
; CURRENT APPLICATION NUMBER: US/09/677,488
; CURRENT FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1263
; TYPE: PRT
; ORGANISM: Euphiala spinifera
; FEATURE:
; OTHER INFORMATION: p-glycoprotein, translation of fully spliced cdna
US-09-677-488-11

```

```

Query Match 100.0%; Score 6453; DB 20; Length 1263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MADESEKPRNQDGSSESHPPPEKETEGSTSDYLIRIFRYADKYDWTNLNIALICAIGSG 60
Db 1 MADESEKPRNQDGSSESHPPPEKETEGSTSDYLIRIFRYADKYDWTNLNIALICAIGSG 60
Qy 61 ASPLMSIIFGFTNKNFNNSGDSPEAFKADVDHFVLPFVYLFICKFVLTIVVSTAAIT 120
Db 61 ASPLMSIIFGFTNKNFNNSGDSPEAFKADVDHFVLPFVYLFICKFVLTIVVSTAAIT 120
Qy 121 ISARTTRTLRRVLECLTQKQVHFDKQSNCAITATVTTNGNRIOGTGIAEKLVTQAL 180
Db 121 ISARTTRTLRRVLECLTQKQVHFDKQSNCAITATVTTNGNRIOGTGIAEKLVTQAL 180
Qy 181 SMFSAFVVALASQWLALITMSVTPAFLVTGICIAIDAAQEARITRYSRAVLAEEV 240
Db 181 SMFSAFVVALASQWLALITMSVTPAFLVTGICIAIDAAQEARITRYSRAVLAEEV 240
Qy 241 LSSIRTVHAFYAKKMYEKYDVLQQAQHEGKKKSPNNNGVLFSTFEYFCIYAALAPFWKG 300
Db 241 LSSIRTVHAFYAKKMYEKYDVLQQAQHEGKKKSPNNNGVLFSTFEYFCIYAALAPFWKG 300
Qy 301 FRMYQNGEVADYGVFTVLSVTLTAATSIISMLAPSGSVVYORRIFGSELSFIIDKPTOLD 360
Db 301 FRMYQNGEVADYGVFTVLSVTLTAATSIISMLAPSGSVVYORRIFGSELSFIIDKPTOLD 360

```

```

Qy 361 PLDSPGQEGCGLGOIIONLAFAYPSRPSAQYLRDFNLTIPAGKTTALVGASGSKSTM 420
Db 361 PLDSPGQEGCGLGOIIONLAFAYPSRPSAQYLRDFNLTIPAGKTTALVGASGSKSTM 420
Qy 421 VGLLERWYLPSSGRILLDGLGQYVNWKLSRIRLVQOEPVLFRTTIFONTANGFMDEQ 480
Db 421 VGLLERWYLPSSGRILLDGLGQYVNWKLSRIRLVQOEPVLFRTTIFONTANGFMDEQ 480
Qy 481 RDLPREKQMLVOKKACASNGDVFINELPNGETEVETGERAGALSGGQORIAIARSIIIS 540
Db 481 RDLPREKQMLVOKKACASNGDVFINELPNGETEVETGERAGALSGGQORIAIARSIIIS 540
Qy 541 PKIILLDEATSDALDQSERLVQA 600
Db 541 PKIILLDEATSDALDQSERLVQA 600
Qy 601 OGTTHIELIEFGCHYAALVRAODLGADQEQHEKTLHEKAAREAGERPALERTHTTATSQ 660
Db 601 OGTTHIELIEFGCHYAALVRAODLGADQEQHEKTLHEKAAREAGERPALERTHTTATSQ 660
Qy 661 AGDLEKRVVPGTGLYSLLKILIMEYEQNLVWCFLLSTITVILICAATFFGQALLFSRL 720
Db 661 AGDLEKRVVPGTGLYSLLKILIMEYEQNLVWCFLLSTITVILICAATFFGQALLFSRL 720
Qy 721 LTVFELSGHAAQERADFYILMFFVVALGNLVGYFTIGMTCNVISOVTVHRYQAAMFORVL 780
Db 721 LTVFELSGHAAQERADFYILMFFVVALGNLVGYFTIGMTCNVISOVTVHRYQAAMFORVL 780
Qy 781 DODIELDIPQISGALTSALPTQLOELISANFLIYVVGQHRLEQCSSTSLMWTG 840
Db 781 DODIELDIPQISGALTSALPTQLOELISANFLIYVVGQHRLEQCSSTSLMWTG 840
Qy 841 PGGCVMTSTPAFCWLPONSSRDEARSKLGLCKRCKACKSSYRDPDRLLIEDSRPACS 900
Db 841 PGGCVMTSTPAFCWLPONSSRDEARSKLGLCKRCKACKSSYRDPDRLLIEDSRPACS 900
Qy 901 PTVLGHVQGLAKTIOSEFWFGRCFGHLSQSMEFLAIALGFCIAVDNMLQVSTTQLANPIS 960
Db 901 PTVLGHVQGLAKTIOSEFWFGRCFGHLSQSMEFLAIALGFCIAVDNMLQVSTTQLANPIS 960
Qy 961 SSWAFCLPVQAAQYLAISTFSTKARSAANYILMLRTLKPTIRETEENKKKGPVGGCPVD 1020
Db 961 SSWAFCLPVQAAQYLAISTFSTKARSAANYILMLRTLKPTIRETEENKKKGPVGGCPVD 1020
Qy 1021 LEDIEFRYRDRSARVLRGVSMTEIPGQFVAVYGASGCKSTLIALSERFYDPTSGRISF 1080
Db 1021 LEDIEFRYRDRSARVLRGVSMTEIPGQFVAVYGASGCKSTLIALSERFYDPTSGRISF 1080
Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPDLYOGSVRENTVLAELAESELCQGLPAPMLWI 1140
Db 1081 AHENIAEMSPRLYRGHMSLVQOEPDLYOGSVRENTVLAELAESELCQGLPAPMLWI 1140
Qy 1141 LSSLYQAKRLAAQGMQFSGGQORIAIARALIRNPKLLLDDEATSDALDQSERLVQA 1200
Db 1141 LSSLYQAKRLAAQGMQFSGGQORIAIARALIRNPKLLLDDEATSDALDQSERLVQA 1200
Qy 1201 ALDEASTSRTTIAVAHRLSTIRNVDVIFVFANGRIAEFTGTHAELQRLGRYEMCLAQSL 1260
Db 1201 ALDEASTSRTTIAVAHRLSTIRNVDVIFVFANGRIAEFTGTHAELQRLGRYEMCLAQSL 1260
Qy 1261 DQA 1263
Db 1261 DQA 1263

```

```

RESULT 5
US-09-677-488A-11
; Sequence 11, Application US/09677488A
; GENERAL INFORMATION:
; APPLICANT: Duwick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob

```

APPLICANT: Folkerts, Otto  
TITLE OF INVENTION: Crasta, Oswald R.  
TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification  
FILE REFERENCE: 35718/204100  
CURRENT APPLICATION NUMBER: US/09/677,488A  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 09/351,224  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 1263  
TYPE: PRT  
ORGANISM: Exophiala spinifera  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 157  
OTHER INFORMATION: Xaa - Any Amino Acid  
US-09-677-488A-11

Query Match 100.0%; Score 6453; DB 20; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MADESEKPRPDGSGESSHPPPEKETEGSISDYLRIFRYADKYDWTNLNLTALICAGSG 60  
Db 1 MADESEKPRPDGSGESSHPPPEKETEGSISDYLRIFRYADKYDWTNLNLTALICAGSG 60  
Qy 61 ASPLMSIIIFGFTKFNKFNKNSGDSPEAFKADYDFHFWFLYIGFVLYVSTAAIT 120  
Db 61 ASPLMSIIIFGFTKFNKFNKNSGDSPEAFKADYDFHFWFLYIGFVLYVSTAAIT 120  
Qy 121 ISAIRTRTLRRVFELECTLRQEVHFDKQNGALATXVTNGNRITQTGAELKFTVQAL 180  
Db 121 ISAIRTRTLRRVFELECTLRQEVHFDKQNGALATXVTNGNRITQTGAELKFTVQAL 180  
Qy 181 SMFSAFVVALASQKALITMSVIPAFIVTGTICIDAADAEARITRIYRAAVLAEEV 240  
Db 181 SMFSAFVVALASQKALITMSVIPAFIVTGTICIDAADAEARITRIYRAAVLAEEV 240  
Qy 241 LSSIRVHAFYAKKWKVKYDVFLOQAHOEGKKSPNNGVLFSTFECIYAAIALAPWK 300  
Db 241 LSSIRVHAFYAKKWKVKYDVFLOQAHOEGKKSPNNGVLFSTFECIYAAIALAPWK 300  
Qy 301 FRMYQGEVADVGVFTVYVTLAATSIWLPAGSVVYQRRIGSELSFIIDKPTOLD 360  
Db 301 FRMYQGEVADVGVFTVYVTLAATSIWLPAGSVVYQRRIGSELSFIIDKPTOLD 360  
Qy 361 PLDPGKQPEGGLGQIEONLAFAYPSRPSAQVLRDFNLITIPAGKTTALVGASGSGKSTM 420  
Db 361 PLDPGKQPEGGLGQIEONLAFAYPSRPSAQVLRDFNLITIPAGKTTALVGASGSGKSTM 420  
Qy 421 VGLLEWYLPSSGRILLDGLGQYVNVKLSRILRVQOEPVLFRTGTFONTANGFMDQ 480  
Db 421 VGLLEWYLPSSGRILLDGLGQYVNVKLSRILRVQOEPVLFRTGTFONTANGFMDQ 480  
Qy 481 RDLPRKQKELVOKKAKSNGDVFINELPNCYETEVEGERAGALSGGQRRIAIRSIIS 540  
Db 481 RDLPRKQKELVOKKAKSNGDVFINELPNCYETEVEGERAGALSGGQRRIAIRSIIS 540  
Qy 541 PKILLDEATSDALDPAEKVQVQALNRYSKDRUTTLTAHKLATVRSAGNIAVISQKIVE 600  
Db 541 PKILLDEATSDALDPAEKVQVQALNRYSKDRUTTLTAHKLATVRSAGNIAVISQKIVE 600  
Qy 601 QGTHHELIEFGCHYAALVRAODLGADEQOHEKTLHEKAAREAAAGRPALERTHTTATSQ 660  
Db 601 QGTHHELIEFGCHYAALVRAODLGADEQOHEKTLHEKAAREAAAGRPALERTHTTATSQ 660  
Qy 661 AGDLEKRPVGTGLYSLLKCLIMFYEQKNLYWCFLSTITVILICATFPQOALLFSRL 720  
Db 661 AGDLEKRPVGTGLYSLLKCLIMFYEQKNLYWCFLSTITVILICATFPQOALLFSRL 720

Qy 721 LTVPELSGHAABERADFYILMFFVVALGNLVGYETIGWTCNVLSOVVTHRYQAAAFQRL 780  
Db 721 LTVPELSGHAABERADFYILMFFVVALGNLVGYETIGWTCNVLSOVVTHRYQAAAFQRL 780  
Qy 781 DDQIELDIPQISGALTSQISALPTQLOELISANFLIYIVGQHRLEOCSTTSWMETG 840  
Db 781 DDQIELDIPQISGALTSQISALPTQLOELISANFLIYIVGQHRLEOCSTTSWMETG 840  
Qy 841 PGGCVWCTSTPAFGWLPQNSRDEARSRLKGLCKRCKWACKRSYRDPDLIFDSRRPCS 900  
Db 841 PGGCVWCTSTPAFGWLPQNSRDEARSRLKGLCKRCKWACKRSYRDPDLIFDSRRPCS 900  
Qy 901 PTVLGHVEOGLAKIIOGFWGRCFGFHLQSMEFLALALGFCIAVDNMLQVSTTOLNFI 960  
Db 901 PTVLGHVEOGLAKIIOGFWGRCFGFHLQSMEFLALALGFCIAVDNMLQVSTTOLNFI 960  
Qy 961 SSWAFCLPYOAAOYLAYSTSETKARSANVILMLRTKPTIRETEENKKKPGVGCPCVD 1020  
Db 961 SSWAFCLPYOAAOYLAYSTSETKARSANVILMLRTKPTIRETEENKKKPGVGCPCVD 1020  
Qy 1021 LEDIEFRRORDSARVLRGVSMTIEPGQFVAYVGAAGCGKSTLIALSERPYDTSGRISF 1080  
Db 1021 LEDIEFRRORDSARVLRGVSMTIEPGQFVAYVGAAGCGKSTLIALSERPYDTSGRISF 1080  
Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYQSVRENVTLALEAESEELCOGRLPARPMLWI 1140  
Db 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYQSVRENVTLALEAESEELCOGRLPARPMLWI 1140  
Qy 1141 LSSYOKALKRLAORGMQFSGORAIARALIRNPKLLLDLDEATSAIDTQSERLVA 1200  
Db 1141 LSSYOKALKRLAORGMQFSGORAIARALIRNPKLLLDLDEATSAIDTQSERLVA 1200  
Qy 1201 ALDEASTSPTTIAVAHRLSTIRNDVIVFVANGRIATGTTHAELOKRGYEMCLAQSL 1260  
Db 1201 ALDEASTSPTTIAVAHRLSTIRNDVIVFVANGRIATGTTHAELOKRGYEMCLAQSL 1260  
Qy 1261 DQA 1263  
Db 1261 DQA 1263

RESULT 6  
US-09-677-682-11  
; Sequence 11, Application US/09677682  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification  
; FILE REFERENCE: 35718/204101 (5718-111B )  
; CURRENT APPLICATION NUMBER: US/09/677,682  
; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1263  
; TYPE: PRT  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: p-glycoprotein, translation of fully spliced cDNA  
US-09-677-682-11

Query Match 100.0%; Score 6453; DB 20; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MADESEKPRPDGSGESSHPPPEKETEGSISDYLRIFRYADKYDWTNLNLTALICAGSG 60  
Db 1 MADESEKPRPDGSGESSHPPPEKETEGSISDYLRIFRYADKYDWTNLNLTALICAGSG 60  
Qy 61 ASPLMSIIIFGFTKFNKFNKNSGDSPEAFKADYDFHFWFLYIGFVLYVSTAAIT 120

Db 61 ASPLMSIIFGFTNKNFNNSGDSPEAFKADVDHFLVFLVIGFVLTYYSTAAT 120  
Qy 121 ISAITRTTTLRRVLECTLRQEVNHFHDKOSNGAIATXVTNGNRITQGTIAEKLVFTVOAL 180  
Db 121 ISAITRTTTLRRVLECTLRQEVNHFHDKOSNGAIATXVTNGNRITQGTIAEKLVFTVOAL 180  
Qy 181 SMFSAFVVALASQWKALITMSVIPAIFLVGTGICIAIDAAQEARITRIYSRAAVLAEV 240  
Db 181 SMFSAFVVALASQWKALITMSVIPAIFLVGTGICIAIDAAQEARITRIYSRAAVLAEV 240  
Qy 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNKPNNGVLFSTFYFCIYAAIALAFWK 300  
Db 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNKPNNGVLFSTFYFCIYAAIALAFWK 300  
Qy 301 FRMYONGEADVGVFTVVLVSVTLAATSI SMLAPSGSVVYQRRIFGSELSFIIDKPTOLD 360  
Db 301 FRMYONGEADVGVFTVVLVSVTLAATSI SMLAPSGSVVYQRRIFGSELSFIIDKPTOLD 360  
Qy 361 PLDPGKQPEGCLQOIEIQNLAFAYPSRPSAQVLRDNLTIPAGKTTALVGASGSGSTM 420  
Db 361 PLDPGKQPEGCLQOIEIQNLAFAYPSRPSAQVLRDNLTIPAGKTTALVGASGSGSTM 420  
Qy 421 VGLLWYLPSSGRILLDGLLEQYVNWLRSRIRLVQOEPVLFRTTIFONTANGFDEQ 480  
Db 421 VGLLWYLPSSGRILLDGLLEQYVNWLRSRIRLVQOEPVLFRTTIFONTANGFDEQ 480  
Qy 481 RDLPREKOMELVQACASNGDVFINELPNGYETEVGERAGALSGQRORAIARSIIISD 540  
Db 481 RDLPREKOMELVQACASNGDVFINELPNGYETEVGERAGALSGQRORAIARSIIISD 540  
Qy 541 PKILLDDATSAIDPAKVKVQALNRVSKDRTLVIAHKLATVKSAGNIIVISOGKIVE 600  
Db 541 PKILLDDATSAIDPAKVKVQALNRVSKDRTLVIAHKLATVKSAGNIIVISOGKIVE 600  
Qy 601 QGTHHELFEFCHYAALVRAODLGADQEQHEKTLHEKAAREAPALERTHTTATSQ 660  
Db 601 QGTHHELFEFCHYAALVRAODLGADQEQHEKTLHEKAAREAPALERTHTTATSQ 660  
Qy 661 AGDLEKRPVPGTLYGSLKILIMFEYQKNLYWCFLSTIVILICAATFPGQALLFSRL 720  
Db 661 AGDLEKRPVPGTLYGSLKILIMFEYQKNLYWCFLSTIVILICAATFPGQALLFSRL 720  
Qy 721 LTVFELSGHAAQERADFYILMEFVVALGNLVGYFTIGWTCNVISOVTHRYQAAAFORVL 780  
Db 721 LTVFELSGHAAQERADFYILMEFVVALGNLVGYFTIGWTCNVISOVTHRYQAAAFORVL 780  
Qy 781 DQDIELDIPQISGALTSQISALPTQLOELISANFLIYVVGQHLRLOCSTTSLWETG 840  
Db 781 DQDIELDIPQISGALTSQISALPTQLOELISANFLIYVVGQHLRLOCSTTSLWETG 840  
Qy 841 PGGCVWCTSTPAFCWLPONSRRDEARSLKGLCKRCWKACKRSSYRDPDRLIFDSRRPCS 900  
Db 841 PGGCVWCTSTPAFCWLPONSRRDEARSLKGLCKRCWKACKRSSYRDPDRLIFDSRRPCS 900  
Qy 901 PTVLGHVQGLAKIQSFWEGRGCFGRHLSQSMFEFLAIALGFCIAVDNMLQVSTTQLNFI 960  
Db 901 PTVLGHVQGLAKIQSFWEGRGCFGRHLSQSMFEFLAIALGFCIAVDNMLQVSTTQLNFI 960  
Qy 961 SSWAFCLPQAAAYLAYSSTFKARSAANYILWLRTKPTIRETEENKKGPGVGGCPVD 1020  
Db 961 SSWAFCLPQAAAYLAYSSTFKARSAANYILWLRTKPTIRETEENKKGPGVGGCPVD 1020  
Qy 1021 LEDIEFRYQRDSARVLRGVSMITEPCQFVAVVGASGCKSTPLIALSERFYDPTSGRISF 1080  
Db 1021 LEDIEFRYQRDSARVLRGVSMITEPCQFVAVVGASGCKSTPLIALSERFYDPTSGRISF 1080  
Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYOGSVRENTVLAELSELSELCQRLPARPMLWI 1140  
Db 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYOGSVRENTVLAELSELSELCQRLPARPMLWI 1140  
Qy 1141 LSSLYOKALKRLAAQRMQFSGQRORIAARALIRNPKILLDDATSAIDTQSERLVOA 1200

Db 1141 LSSLYOKALKRLAAQRMQFSGQRORIAARALIRNPKILLDDATSAIDTQSERLVOA 1200  
Qy 1201 ALDEASTSRRTIAVAHRLSTIRNVDFVFPANGRIAEFTGTHAELORLGRYYEMCLAQSL 1260  
Db 1201 ALDEASTSRRTIAVAHRLSTIRNVDFVFPANGRIAEFTGTHAELORLGRYYEMCLAQSL 1260  
Qy 1261 DQA 1263  
Db 1261 DQA 1263  
RESULT 7  
US-09-677-682A-11  
: Sequence 11, Application US/09677682A  
: GENERAL INFORMATION:  
: APPLICANT: Duwick, Jon  
: APPLICANT: Maddox, Joyce  
: APPLICANT: Gilliam, Jacob  
: APPLICANT: Folkerts, Otto  
: APPLICANT: Crasta, Oswald R.  
: TITLE OF INVENTION: Compositions and Methods for Fumonisin  
: FILE REFERENCE: 35718/204101  
: CURRENT APPLICATION NUMBER: US/09/677,682A  
: PRIOR FILING DATE: 2000-10-02  
: PRIOR APPLICATION NUMBER: 09/351,224  
: PRIOR FILING DATE: 1999-07-12  
: NUMBER OF SEQ ID NOS: 11  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 11  
: LENGTH: 1263  
: TYPE: PRT  
: ORGANISM: Exophiala spinifera  
: FEATURE:  
: NAME/KEY: VARIANT  
: LOCATION: 157  
: OTHER INFORMATION: Xaa - Any Amino Acid  
US-09-677-682A-11  
Query Match 100.0%; Score 6453; DB 20; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MADESEKPRNPDGSESSHPPEKEGTSIDYLRIFRYADKYDWTLVIALICAISG 60  
Db 1 MADESEKPRNPDGSESSHPPEKEGTSIDYLRIFRYADKYDWTLVIALICAISG 60  
Qy 61 ASPLMSIIFGFTNKNFNNSGDSPEAFKADVDHFLVFLVIGFVLTYYSTAAT 120  
Db 61 ASPLMSIIFGFTNKNFNNSGDSPEAFKADVDHFLVFLVIGFVLTYYSTAAT 120  
Qy 121 ISAITRTTTLRRVLECTLRQEVNHFHDKOSNGAIATXVTNGNRITQGTIAEKLVFTVOAL 180  
Db 121 ISAITRTTTLRRVLECTLRQEVNHFHDKOSNGAIATXVTNGNRITQGTIAEKLVFTVOAL 180  
Qy 181 SMFSAFVVALASQWKALITMSVIPAIFLVGTGICIAIDAAQEARITRIYSRAAVLAEV 240  
Db 181 SMFSAFVVALASQWKALITMSVIPAIFLVGTGICIAIDAAQEARITRIYSRAAVLAEV 240  
Qy 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNKPNNGVLFSTFYFCIYAAIALAFWK 300  
Db 241 LSSIRTVHAFYAQKKWVEKYDVFLOQAHOEGKKSPNNKPNNGVLFSTFYFCIYAAIALAFWK 300  
Qy 301 FRMYONGEADVGVFTVVLVSVTLAATSI SMLAPSGSVVYQRRIFGSELSFIIDKPTOLD 360  
Db 301 FRMYONGEADVGVFTVVLVSVTLAATSI SMLAPSGSVVYQRRIFGSELSFIIDKPTOLD 360  
Qy 361 PLDPGKQPEGCLQOIEIQNLAFAYPSRPSAQVLRDNLTIPAGKTTALVGASGSGSTM 420  
Db 361 PLDPGKQPEGCLQOIEIQNLAFAYPSRPSAQVLRDNLTIPAGKTTALVGASGSGSTM 420  
Qy 421 VGLLWYLPSSGRILLDGLLEQYVNWLRSRIRLVQOEPVLFRTTIFONTANGFDEQ 480

Db 421 VGLLEWYLPSSGRILLDGLGQYNNKWLRSIRLVQEPVLFRTIFQNIANGFMDEQ 480  
Qy 481 RDLPRKQELVQKACKASNGDVFINELPNGYETEVGERAGALSGGQRIARSIISD 540  
Db 481 RDLPRKQELVQKACKASNGDVFINELPNGYETEVGERAGALSGGQRIARSIISD 540  
Qy 541 PKILLDEATSALDPKAEKVYQVEALNRVSKDRITLVIAHKLATVKSAGNIAVISOGKIVE 600  
Db 541 PKILLDEATSALDPKAEKVYQVEALNRVSKDRITLVIAHKLATVKSAGNIAVISOGKIVE 600  
Qy 601 QGTHHELIEFGCHYAALVRAODLGADQOEHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
Db 601 QGTHHELIEFGCHYAALVRAODLGADQOEHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
Qy 661 AGDLEKRPVPGTGLYSLLKCLIMFYEQKNLYWCFLSTITVILICAATFFQCALFSRL 720  
Db 661 AGDLEKRPVPGTGLYSLLKCLIMFYEQKNLYWCFLSTITVILICAATFFQCALFSRL 720  
Qy 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAAMFORVL 780  
Db 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAAMFORVL 780  
Qy 781 DODIELDIPQISGALTSOLSALPTQELISANFLIYVVGQHLRBOCSTTSLWMTG 840  
Db 781 DODIELDIPQISGALTSOLSALPTQELISANFLIYVVGQHLRBOCSTTSLWMTG 840  
Qy 841 PGCVWCTSTPAFGWLPONSSRDEARSKLGLCKRCWACKRSSYRDPDLIEDSRPCS 900  
Db 841 PGCVWCTSTPAFGWLPONSSRDEARSKLGLCKRCWACKRSSYRDPDLIEDSRPCS 900  
Qy 901 PTVIGHVEGLAKIIQSFWRCFCFHLSSQSMFEFLAIALGFCIAVDNNLQVSTTOLNFI 960  
Db 901 PTVIGHVEGLAKIIQSFWRCFCFHLSSQSMFEFLAIALGFCIAVDNNLQVSTTOLNFI 960  
Qy 961 SSWAFCLPQAAQAYLAYSTSTFKARSAANYILWLRLKPTIRETEENKKGPGVGGCPVD 1020  
Db 961 SSWAFCLPQAAQAYLAYSTSTFKARSAANYILWLRLKPTIRETEENKKGPGVGGCPVD 1020  
Qy 1021 LEDIEFRYRDSARVLGVSMTLEPGOFVAYVYGASGCKSTLIALSERFYDPTSGRISF 1080  
Db 1021 LEDIEFRYRDSARVLGVSMTLEPGOFVAYVYGASGCKSTLIALSERFYDPTSGRISF 1080  
Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYGGSVRENVTLEAEGLSELCQGRLPARPLMI 1140  
Db 1081 AHENIAEMSPRLYRGHMSLVQOEPVLYGGSVRENVTLEAEGLSELCQGRLPARPLMI 1140  
Qy 1141 LSSLYOKALKRLAARGMFGSGGQRIARALINPKLLLLDEATSALDTSQSERLVQA 1200  
Db 1141 LSSLYOKALKRLAARGMFGSGGQRIARALINPKLLLLDEATSALDTSQSERLVQA 1200  
Qy 1201 ALDEASTSTTIAVAHRLSTIRNVDVIFVFPANGRIAEFTGTHAELQRLGRYVYEMCLAQSL 1260  
Db 1201 ALDEASTSTTIAVAHRLSTIRNVDVIFVFPANGRIAEFTGTHAELQRLGRYVYEMCLAQSL 1260  
Qy 1261 DOA 1263  
Db 1261 DOA 1263

RESULT 8  
us-09-677-682b-11  
; Sequence 11, Application US/09677682B  
; GENERAL INFORMATION:  
; APPLICANT: Duwick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE OF INVENTION: Detoxification  
; FILE REFERENCE: 35718/204101  
; CURRENT APPLICATION NUMBER: US/09/677,682B

; CURRENT FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1263  
; TYPE: PRT  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 157  
; OTHER INFORMATION: xaa - Any Amino Acid  
US-09-677-682b-11  
  
Query Match 100.00; Score 6453; DB 20; Length 1263;  
Best local similarity 100.00; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MADESEKPRPNQDGSSESSHPPEKETEGSISDYLIRIFRYADKYDWTNLVIALICAISG 60  
Db 1 MADESEKPRPNQDGSSESSHPPEKETEGSISDYLIRIFRYADKYDWTNLVIALICAISG 60  
Qy 61 ASPLPLMSIIFGSFTNKNFNNSGDSPEAFKADVDHFWLVYLFYFGKFLVYVSTAAT 120  
Db 61 ASPLPLMSIIFGSFTNKNFNNSGDSPEAFKADVDHFWLVYLFYFGKFLVYVSTAAT 120  
Qy 121 ISAIRTRRLRRVFECLTQEVWHFQKNGALATAYVTNGHRIOTGIAEKLVTQAL 180  
Db 121 ISAIRTRRLRRVFECLTQEVWHFQKNGALATAYVTNGHRIOTGIAEKLVTQAL 180  
Qy 181 SMFSPAPVVALASOMKALITMSVIPAIPLVTGICIAIDAAQARITRIYSRAAVLAEV 240  
Db 181 SMFSPAPVVALASOMKALITMSVIPAIPLVTGICIAIDAAQARITRIYSRAAVLAEV 240  
Qy 241 LSSIRTVHAFYAKKMKERYDVFYQQAHQEGKKSPNNGVLFSTEFYFCIYAAIALAFWK 300  
Db 241 LSSIRTVHAFYAKKMKERYDVFYQQAHQEGKKSPNNGVLFSTEFYFCIYAAIALAFWK 300  
Qy 301 FRWYQGEVADVKVFTVSVTLAATSISMLAPSGSVYVORRIFGSELSIIDKPTOLD 360  
Db 301 FRWYQGEVADVKVFTVSVTLAATSISMLAPSGSVYVORRIFGSELSIIDKPTOLD 360  
Qy 361 PLDPKQEGCGCQIEQNLAFAPYSPSAOVLNDFNLTPAGKTTALVAGSGSKSTM 420  
Db 361 PLDPKQEGCGCQIEQNLAFAPYSPSAOVLNDFNLTPAGKTTALVAGSGSKSTM 420  
Qy 421 VGLLEWYLPSSGRILLDGLGQYNNKWLRSIRLVQEPVLFRTIFQNIANGFMDEQ 480  
Db 421 VGLLEWYLPSSGRILLDGLGQYNNKWLRSIRLVQEPVLFRTIFQNIANGFMDEQ 480  
Qy 481 RDLPRKQELVQKACKASNGDVFINELPNGYETEVGERAGALSGGQRIARSIISD 540  
Db 481 RDLPRKQELVQKACKASNGDVFINELPNGYETEVGERAGALSGGQRIARSIISD 540  
Qy 541 PKILLDEATSALDPKAEKVYQVEALNRVSKDRITLVIAHKLATVKSAGNIAVISOGKIVE 600  
Db 541 PKILLDEATSALDPKAEKVYQVEALNRVSKDRITLVIAHKLATVKSAGNIAVISOGKIVE 600  
Qy 601 QGTHHELIEFGCHYAALVRAODLGADQOEHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
Db 601 QGTHHELIEFGCHYAALVRAODLGADQOEHEKTLHEKAAREAAAGERPALERTHTTATSQ 660  
Qy 661 AGDLEKRPVPGTGLYSLLKCLIMFYEQKNLYWCFLSTITVILICAATFFQCALFSRL 720  
Db 661 AGDLEKRPVPGTGLYSLLKCLIMFYEQKNLYWCFLSTITVILICAATFFQCALFSRL 720  
Qy 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAAMFORVL 780  
Db 721 LTVFELSGHAAQERADFYILMFVVALGNLVGYFTIGWTCNVISQVTHRYQAAMFORVL 780  
Qy 781 DODIELDIPQISGALTSOLSALPTQELISANFLIYVVGQHLRBOCSTTSLWMTG 840

Db 781 DODIELDIPROISGALTSQSLPQLOELISANFLIYVVGQRLBQCSTTSLWMTG 840  
 Qy 841 PGGCVWCTSTPAFCWLPONSSRDEARSRKLGKLCRCWACKRSSYRDPDRILFDSRRPCS 900  
 Db 841 PGGCVWCTSTPAFCWLPONSSRDEARSRKLGKLCRCWACKRSSYRDPDRILFDSRRPCS 900  
 Qy 901 PTVLGHVEOGLAKIIQSFWFGRCFGPHLSQSMFLAIALGFCIAVDNWLQVSTTQLNFI 960  
 Db 901 PTVLGHVEOGLAKIIQSFWFGRCFGPHLSQSMFLAIALGFCIAVDNWLQVSTTQLNFI 960  
 Qy 961 SSWAFCLPVQAAQYLAYSFTFKARSAANYILWLRTLKPTIRETEENKKKPGVGGCPVD 1020  
 Db 961 SSWAFCLPVQAAQYLAYSFTFKARSAANYILWLRTLKPTIRETEENKKKPGVGGCPVD 1020  
 Qy 1021 LEDIEFRYRQDSARVLRGVSMTEPCQFVAVYGSGCGKSTLIALSERFYDPTSGRISF 1080  
 Db 1021 LEDIEFRYRQDSARVLRGVSMTEPCQFVAVYGSGCGKSTLIALSERFYDPTSGRISF 1080  
 Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPITYOGSVRENTVLAELAESELCQGRPARPLMI 1140  
 Db 1081 AHENIAEMSPRLYRGHMSLVQOEPITYOGSVRENTVLAELAESELCQGRPARPLMI 1140  
 Qy 1141 LSSUYOKALKRAAQRMQSGGQORQRIARALIRNPKLLLDLDEATSDALDQSERLVQA 1200  
 Db 1141 LSSUYOKALKRAAQRMQSGGQORQRIARALIRNPKLLLDLDEATSDALDQSERLVQA 1200  
 Qy 1201 ALDEASTSRTTIAVAHRLSTIRNVDFVIFVANGRIAGTGAELQRLGRYYEMCLAQSL 1260  
 Db 1201 ALDEASTSRTTIAVAHRLSTIRNVDFVIFVANGRIAGTGAELQRLGRYYEMCLAQSL 1260  
 Qy 1261 DQA 1263  
 Db 1261 DQA 1263

RESULT 9

US-09-882-694-11  
 ; Sequence 11, Application US/09882694  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Duwick, Jon  
 ; APPLICANT: Maddox, Joyce  
 ; APPLICANT: Gilliam, Jacob  
 ; APPLICANT: Folkerts, Otto  
 ; APPLICANT: Crasta, Oswald R.  
 ; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
 ; FILE REFERENCE: 35718/208255  
 ; CURRENT APPLICATION NUMBER: US/09/882,694  
 ; CURRENT FILING DATE: 2001-06-12  
 ; PRIOR APPLICATION NUMBER: 09/351,224  
 ; PRIOR FILING DATE: 1999-07-12  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11  
 ; LENGTH: 1263  
 ; TYPE: PRT  
 ; ORGANISM: Exophiala spinifera  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: 157  
 ; OTHER INFORMATION: Xaa - Any Amino Acid  
 US-09-882-694-11

Query Match 100.0%; Score 6453; DB 22; Length 1263;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MADESEKPRPNQDSESSSSHPPEKETEGSIDYLRIFRYADKYDWTNLNIALICATGSG 60  
 Db 1 MADESEKPRPNQDSESSSSHPPEKETEGSIDYLRIFRYADKYDWTNLNIALICATGSG 60

Qy 61 ASPLMSIIFGSETNKNFNYSNGDGSPEAFKADVDHFWLWVYLFYIGKFLVYVYSTAIT 120  
 Db 61 ASPLMSIIFGSETNKNFNYSNGDGSPEAFKADVDHFWLWVYLFYIGKFLVYVYSTAIT 120  
 Qy 121 ISARTTTRTLRRVFLPCTLRQEVWHPDKOSNGAIATXVTTCNRIOTGIAEKLVFTVQAL 180  
 Db 121 ISARTTTRTLRRVFLPCTLRQEVWHPDKOSNGAIATXVTTCNRIOTGIAEKLVFTVQAL 180  
 Qy 181 SMFSAFVVALASQWKLALITHSVIPALPLVTGICIAIDAAQEARITRIYSRAVLAEEV 240  
 Db 181 SMFSAFVVALASQWKLALITHSVIPALPLVTGICIAIDAAQEARITRIYSRAVLAEEV 240  
 Qy 241 LSSIRTYAHFYAOKKWKVEKYDVLQQAHOEGKKSPNNGVLFSTFECYIAAIALAFWKG 300  
 Db 241 LSSIRTYAHFYAOKKWKVEKYDVLQQAHOEGKKSPNNGVLFSTFECYIAAIALAFWKG 300  
 Qy 301 FRMYONGEADVGVFTVVLSTLAATSISMLAPSGSVVYQRRIFGSELSIIOKPTOLD 360  
 Db 301 FRMYONGEADVGVFTVVLSTLAATSISMLAPSGSVVYQRRIFGSELSIIOKPTOLD 360  
 Qy 361 PLDPGKOPCECLGOIEIONLAFAYPSRPSAOVLROFNLTIPACKTTALVCSGSGKSTM 420  
 Db 361 PLDPGKOPCECLGOIEIONLAFAYPSRPSAOVLROFNLTIPACKTTALVCSGSGKSTM 420  
 Qy 421 VGLLEWYLFSSGRILLDGLGQYNNKWLRSIRLVQOEPVLPFGTIFQNIANGFMDQ 480  
 Db 421 VGLLEWYLFSSGRILLDGLGQYNNKWLRSIRLVQOEPVLPFGTIFQNIANGFMDQ 480  
 Qy 481 RDLPREKQOMELVQKACKASNGDVFINELPNGYETEVGERAGALSQGRQRIARSIISD 540  
 Db 481 RDLPREKQOMELVQKACKASNGDVFINELPNGYETEVGERAGALSQGRQRIARSIISD 540  
 Qy 541 PKILLDEATSDALDPAEKVVOEALNRVSKDRITLVIAHKLATYKSNAGNIAVISOQKIVE 600  
 Db 541 PKILLDEATSDALDPAEKVVOEALNRVSKDRITLVIAHKLATYKSNAGNIAVISOQKIVE 600  
 Qy 601 QGTHHELIFEGCHYALVRAQDLGADGQEHKTLHEKAAREMAGERPALERTHTTTSQ 660  
 Db 601 QGTHHELIFEGCHYALVRAQDLGADGQEHKTLHEKAAREMAGERPALERTHTTTSQ 660  
 Qy 661 AGDLEKRPVPGTGLYSLKCLILIMFYEQNLWYCFLLSTITVLICAAATFPQALLFSRL 720  
 Db 661 AGDLEKRPVPGTGLYSLKCLILIMFYEQNLWYCFLLSTITVLICAAATFPQALLFSRL 720  
 Qy 721 LTVFELSGHAAQERADFYILMPFVVALGNLVGYFTIGWTCNVISQVTHRYQAAAMFORVL 780  
 Db 721 LTVFELSGHAAQERADFYILMPFVVALGNLVGYFTIGWTCNVISQVTHRYQAAAMFORVL 780  
 Qy 781 DODIELLDIPEOISGALTSQSLPQLOELISANFLIYVVGQRLBQCSTTSLWMTG 840  
 Db 781 DODIELLDIPEOISGALTSQSLPQLOELISANFLIYVVGQRLBQCSTTSLWMTG 840  
 Qy 841 PGGCVWCTSTPAFCWLPONSSRDEARSRKLGKLCRCWACKRSSYRDPDRILFDSRRPCS 900  
 Db 841 PGGCVWCTSTPAFCWLPONSSRDEARSRKLGKLCRCWACKRSSYRDPDRILFDSRRPCS 900  
 Qy 901 PTVLGHVEOGLAKIIQSFWFGRCFGPHLSQSMFLAIALGFCIAVDNWLQVSTTQLNFI 960  
 Db 901 PTVLGHVEOGLAKIIQSFWFGRCFGPHLSQSMFLAIALGFCIAVDNWLQVSTTQLNFI 960  
 Qy 961 SSWAFCLPVQAAQYLAYSFTFKARSAANYILWLRTLKPTIRETEENKKKPGVGGCPVD 1020  
 Db 961 SSWAFCLPVQAAQYLAYSFTFKARSAANYILWLRTLKPTIRETEENKKKPGVGGCPVD 1020  
 Qy 1021 LEDIEFRYRQDSARVLRGVSMTEPCQFVAVYGSGCGKSTLIALSERFYDPTSGRISF 1080  
 Db 1021 LEDIEFRYRQDSARVLRGVSMTEPCQFVAVYGSGCGKSTLIALSERFYDPTSGRISF 1080  
 Qy 1081 AHENIAEMSPRLYRGHMSLVQOEPITYOGSVRENTVLAELAESELCQGRPARPLMI 1140  
 Db 1081 AHENIAEMSPRLYRGHMSLVQOEPITYOGSVRENTVLAELAESELCQGRPARPLMI 1140  
 Qy 1141 LSSUYOKALKRAAQRMQSGGQORQRIARALIRNPKLLLDLDEATSDALDQSERLVQA 1200

Db 1141 LSSLYQKALKRLAAGRGWFGSGGQRRIATARALINPKLLLLDDEATSDTQSERLVA 1200  
Qy 1201 ALDEASTRTTIAVAHRLSTIRNVDFVFPANGRIAGTGTTHAELQRLGRYEMCLAQSL 1260  
Db 1201 ALDEASTRTTIAVAHRLSTIRNVDFVFPANGRIAGTGTTHAELQRLGRYEMCLAQSL 1260  
Qy 1261 DOA 1263  
Db 1261 DOA 1263  
RESULT 10  
US-09-882-694A-11  
; Sequence 11, Application US/09882694A  
; GENERAL INFORMATION:  
; APPLICANT: Duivick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/882.694A  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: fastseq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1263  
; TYPE: PRT  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 157  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-882-694A-11

Query Match 100.08; Score 6453; DB 22; Length 1263;  
Best Local Similarity 100.08; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MADESEKPRNQDGSSESHPPPEKETEGSISDYLRIFRYADKYDWTLNWIALICAIGSG 60  
Db 1 MADESEKPRNQDGSSESHPPPEKETEGSISDYLRIFRYADKYDWTLNWIALICAIGSG 60  
Qy 61 ASPLPMSIIFGSFTNKNFNNSGDCGSPFAKADVDHFLVFLFTGKFLVTVVSTAAIT 120  
Db 61 ASPLPMSIIFGSFTNKNFNNSGDCGSPFAKADVDHFLVFLFTGKFLVTVVSTAAIT 120  
Qy 121 ISARTTTRLRVFLCTLRQEVHFDKQSNAGATATVTTNGRIQTGAELVFTVOAL 180  
Db 121 ISARTTTRLRVFLCTLRQEVHFDKQSNAGATATVTTNGRIQTGAELVFTVOAL 180  
Qy 181 SMFSAFVVALASQWKLALIMSVIPALFVITGICIAIDAAQEARITRIYRAAVLAEEV 240  
Db 181 SMFSAFVVALASQWKLALIMSVIPALFVITGICIAIDAAQEARITRIYRAAVLAEEV 240  
Qy 241 LSSIRTVHAFYAKKMKVEYIDYFLQQAHOEGKKSPNNGVLFSTFYFCIYAAIALAFWK 300  
Db 241 LSSIRTVHAFYAKKMKVEYIDYFLQQAHOEGKKSPNNGVLFSTFYFCIYAAIALAFWK 300  
Qy 301 FRMYQGEVADVGVFTVSVLTAAISIMLAPSGSVYQRRIFGSELSFIIDKPTQD 360  
Db 301 FRMYQGEVADVGVFTVSVLTAAISIMLAPSGSVYQRRIFGSELSFIIDKPTQD 360  
Qy 361 PLDPGKQPGECGLQTEIONLAFATPSPSAQVLRDNFLIPACKTTALVGAGSGKSTM 420  
Db 361 PLDPGKQPGECGLQTEIONLAFATPSPSAQVLRDNFLIPACKTTALVGAGSGKSTM 420

Qy 421 VGLLERNYLPSSGRILLDGLGOYNVWKLRSIRLVOOEPVLPFRGTIFONIANGFMDQ 480  
Db 421 VGLLERNYLPSSGRILLDGLGOYNVWKLRSIRLVOOEPVLPFRGTIFONIANGFMDQ 480  
Qy 481 RDLPREKOMELVOKAKASNDGVFINELPNGYETEVGERAGALSGGORIATARIISD 540  
Db 481 RDLPREKOMELVOKAKASNDGVFINELPNGYETEVGERAGALSGGORIATARIISD 540  
Qy 541 PKILLDEATSDALDPKAEKVVOBALNRVSKORTTLVIAHKLATVKSAGNIAVISQKIYE 600  
Db 541 PKILLDEATSDALDPKAEKVVOBALNRVSKORTTLVIAHKLATVKSAGNIAVISQKIYE 600  
Qy 601 QGTHHELIERGCHYAALVRAQDLGADGDEQOEHEKTLHEKAAREAAAGERPALERTHTTTSQ 660  
Db 601 QGTHHELIERGCHYAALVRAQDLGADGDEQOEHEKTLHEKAAREAAAGERPALERTHTTTSQ 660  
Qy 661 AGDLEKRVPGVTIGYSLKCLILIMEYEOKNLYWCFLSTITVLICAAATPGQALLFSRL 720  
Db 661 AGDLEKRVPGVTIGYSLKCLILIMEYEOKNLYWCFLSTITVLICAAATPGQALLFSRL 720  
Qy 721 LTVFELSCHAAQERADPYILMPFVVALGNLYVFTIGTNCVVISQVYTHRYQAAAFQVRL 780  
Db 721 LTVFELSCHAAQERADPYILMPFVVALGNLYVFTIGTNCVVISQVYTHRYQAAAFQVRL 780  
Qy 781 DDJIELLDIPEQISGALTSQLSALPTQLQELISANFLIYVVGQHRLEQCSSTTSLAMETG 840  
Db 781 DDJIELLDIPEQISGALTSQLSALPTQLQELISANFLIYVVGQHRLEQCSSTTSLAMETG 840  
Qy 841 PGGCVWCTSTPAFGWLPONSSRDEARSRKGLKCHKCKWACKRSYRDPDRILFDSRRPCS 900  
Db 841 PGGCVWCTSTPAFGWLPONSSRDEARSRKGLKCHKCKWACKRSYRDPDRILFDSRRPCS 900  
Qy 901 PTVLGHVEQGLAKTIQSFWEFCRFGHLSQSMFELAIAGFCIAVDNKLQVSTTOLNFI 960  
Db 901 PTVLGHVEQGLAKTIQSFWEFCRFGHLSQSMFELAIAGFCIAVDNKLQVSTTOLNFI 960  
Qy 961 SSWAFCLPVQAAQYLAYSTFTKARSAANYILWRLTKPTIRETEENKKKPGVGGCPVD 1020  
Db 961 SSWAFCLPVQAAQYLAYSTFTKARSAANYILWRLTKPTIRETEENKKKPGVGGCPVD 1020  
Qy 1021 LEDIEFRYRDSARVLRGVSMTEPGQFVAVVGASGCGKSTLIALSERFYDPTSGRISF 1080  
Db 1021 LEDIEFRYRDSARVLRGVSMTEPGQFVAVVGASGCGKSTLIALSERFYDPTSGRISF 1080  
Qy 1081 AHENTAMSPRLYRGHMSLVQOEPYLYOGSVRENTVLAELSELCOGRLPARPMLWI 1140  
Db 1081 AHENTAMSPRLYRGHMSLVQOEPYLYOGSVRENTVLAELSELCOGRLPARPMLWI 1140  
Qy 1141 LSSLYQKALKRLAAGRGWFGSGGQRRIATARALINPKLLLLDDEATSDTQSERLVA 1200  
Db 1141 LSSLYQKALKRLAAGRGWFGSGGQRRIATARALINPKLLLLDDEATSDTQSERLVA 1200  
Qy 1201 ALDEASTRTTIAVAHRLSTIRNVDFVFPANGRIAGTGTTHAELQRLGRYEMCLAQSL 1260  
Db 1201 ALDEASTRTTIAVAHRLSTIRNVDFVFPANGRIAGTGTTHAELQRLGRYEMCLAQSL 1260  
Qy 1261 DOA 1263  
Db 1261 DOA 1263  
RESULT 11  
PCT-US98-27499-2  
; Sequence 2, Application PC/TUS9827499  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrC of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center





```

> FILING DATE:
> CLASSIFICATION: 435
> ATTORNEY/AGENT INFORMATION:
> NAME: Hamilton, Amy E
> REGISTRATION NUMBER: 33,894
> REFERENCE/DOCKET NUMBER: X-9681
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 317-276-3169
> TELEFAX: 317-276-1294
> INFORMATION FOR SEQ ID NO: 2:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 1349 amino acids
> TYPE: amino acid
> TOPOLOGY: linear
> MOLECULE TYPE: protein
> PS-08-612-734-2

```

Query Match 30.3%; Score 1956; DB 10; Length 1349;  
Best Local Similarity 35.8%; Pred. No. 5.9e-178;  
Matches 491; Conservative 229; Mismatches 478; Indels 172; Gaps 29;

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1334 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US98-27300-2

Query Match 30.18; Score 1940; DB 1: Length 1334;  
Best Local Similarity 36.88; Pred. No. 28-176;  
Matches 490; Conservative 231; Mismatches 483; Indels 128; Gaps 30:  
QY 2 ADSEKPRNQDSESS-----SHPPPE-----KETEGSIDYLRIFRYADKY 44  
DB 54 ADKEHPKSSNNNAVSYNEVDALIAHLPEDEQVILKTQLEBIKYNIS-FFGLWRATKM 112  
QY 45 DMTLVNIALICAGSGASLPLMSIIFGSPNKNYNSGDGSPFAKADVDHFLWFEVL 104  
DB 113 DILMIVISTICAIAASTFORIMLYQISY-----DEFYDELTKNLYFEVL 158  
QY 105 FIGKFLVTYVAAATITSAIATTTTLRRVFECLTRQEVVHFDKOSNGAIAATXVTNGNR 164  
DB 159 GIGEFVTYVSVGVFIYTGHEATQKIREYVLESILRQNGYEDKLGAGEVTTTRITADTNL 218  
QY 165 IQTGIAEKLVTVQALSMFFSAFVVALASOMKIALITMSVIPALVLTGTGICIAIDAAQEA 224  
DB 219 IQDGISEKVGCLTLTALATFTVAFIIAVYKVKLAGLICSSTIVALVLTGSGSOFIYKSK 278  
QY 225 RITRIYRAVLAELVSSITVAFYAKKWEKYDNPLOQAQHQEGKKSPNNGVLFST 284  
DB 279 KSLDSYGACGTVAEEVISIRNATAGTQDKLAKQTEVHLDDEAKGKTKNOIVMGWIGA 338  
QY 285 EYFCIIAATAAFWGFPMYONGEVADYGVVYVLSVTLAATISMLAPSGSVVYQRI 344  
DB 339 MFLMTNGLGFMGWSRFLVDGAV-DVGDLITVLMALIGFSIGNVSPNAQAFNVA 397  
QY 345 FGSELSIIDKTQDLPDQSGKQPECCQIQEIQIMAFYSPRSQAOLVDFNLITPAG 404  
DB 398 AAKIETGIDRQSPIDPYNSNGKTLDFHGEHIELNRVHIYSPREVTVMEDYSLSPAG 457  
QY 405 KTTALVAGSGSKTMVGLERWYLPSSGRILLDGLGQYKYNKLSRIRLVOQEPVLF 464  
DB 458 KTTALVPGSGSKTMVGLVRFYMPVGRGTVLDDGHDIKDLNRLRQOISLVSQEPVLF 517  
QY 465 RGTIFQNIANGFMDQRDLPRE-KQELVQKACKASNGDVFINELPNQYETEYGERAGAL 523  
DB 518 GTTIYKNIHRGLIGTKYENESDKVRELIENAAKMANAHDFITALPEGYETNVGQGFLL 577  
QY 524 SGGQRQRIARSIISDPKILLDEATSAIDPKAEKVVOEALNRVSKDRTPVIAHKLAT 583  
DB 578 SGGOKORIAARAVSOPKILLDEATSAIDTKSEGVQAAALERAAREGRTTIVIAHRLST 637  
QY 584 VKSAGNIAYISGKVEQGTHEHIEFGCHYAALVRAQDLGADQOEHE----- 632  
DB 638 IKTAHIVLVNCKIAEQTGHDVLDVDRGGAYKLVQAQRI--NEQKEADALEDADAEDLT 695  
QY 633 -----KTLHEKAREAGERPALERTHTTATSQAGDLEKXKVPVGTGLYSLLKCI-LI 684  
DB 696 NADIAIKITA-SSASDLDGKPTTIDRTGTHKSVSAILLSRRP-PETTPKYSMTLLKVF 753  
QY 685 MFTQKNLWCFLLSTITVILCAATFPQALLPSRLTVFEL---SGHAQERADFYLM 741  
DB 754 ASPNRPEIDY-MLIGLVSVIAGGQOPTQAVLYAKAISTLSLPESQVSKRHADAFWSLM 812  
QY 742 FFVALGNLVGFTIGWTCNVISQVTHRYQAAAMFORVLDQDIELLOPEIQSALTSOL 801  
DB 813 FFVVGIIQITQSTNGAAFAVCSESLIRARSTAFRTILRQDIAFFQKEENSTGALTSFL 872  
QY 802 SALPTQLQEL-----ISANFLIYIVGQHRLEQC-STTSLNWTGPGGCVW 846

DB 873 STETKHLSGSVGVTGLTILMTSTTLGAAIIIALIGWKLALVCISSVVPVLLACG----- 926  
QY 847 CYSTTAFQMLPQNSSRDEARSRLKGLCKRWACKRRSSYRDPRLIFDSRRPCSPVVLGH 906  
DB 927 ---FYRFTYMLAQFQSR-----KLAYEGSANFACEATS---SIRTVASLTRERDVWEIYH 975  
QY 907 VE---OG---LAKIIOSFWFCRGFHLSSOMELAIAGFCIAVDNWLQVSTQNL--- 957  
DB 976 AQDDAQGRSTLSIVRSLL-----YASSQALVFCVALGP-----W-YGTLGLCHHE 1022  
QY 958 -----FTSSSWAECLPVQAAAYLAYSTSTFKARSAANYILWRLTKLPIRE-TENKK 1010  
DB 1023 YDIFRFFVCFPS-ETLGAQAGAGVFSFADPMGAKAKNAARFRFLDRKPKQIDNWSSEGEK 1081  
QY 1011 KQPVGGCPVLEDTIEFRYRORDSARVLRCVSMTEIPGQFVAVGASCGCKSTTIALSERP 1070  
DB 1082 LETVEG-EIEFRNVHFYRPTPEQPVLRGLDLTVKPGYVALVGPSCGKSTTIALLERF 1140  
QY 1071 YDPTSGRISFAHENIAEMSPRLYRGHMSLVQOEPYLYGQSVRENVTLAL-EAELEE-LC 1128  
DB 1141 YDAIAGSLVDGCKDISKLNINSYRSLVSQEPYLYGQSVRENVTLAL-EAELEE-LC 1200  
QY 1129 OQRLPAREPMLWILSSLYQKALKRLAAQRCQWQFSGQORQRIARALIRNPKLLEDEATS 1188  
DB 1201 KACKDANIYDFIMS--LPEGFNTVVVGKGMGLSGKQKQVRAIARALLRDPKILLDEATS 1258  
QY 1189 ALDTOSERLVAALDEASTSTRTTIAVAHRLSTIRNVDFVIFVFNANGRIAEATGTHAELOLR 1248  
DB 1259 ALDSESEKVVQAAALDANAARGRTTIAVAHRLSTIQKADVIYVFDQKIVESGTHSELVQNK 1318  
QY 1249 GRVYEMKLAOSL 1260  
DB 1319 GRVYELVNLQSL 1330

RESULT 14  
US-08-395-246A-2  
Sequence 2, Application US/08395246A  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
TITLE OF INVENTION: ASPERGILLUS FLAVUS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/395,246A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas G.  
REGISTRATION NUMBER: 35784  
REFERENCE/DOCKET NUMBER: x9683  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
TELEFAX: 317-277-1917  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1307 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-395-246A-2

```
Query Match      27.6%  Score 1780;  DB 7;  Length 1307;
Best Local Similarity 34.3%  Pred. No. 5.6e-161;
Matches 466;  Conservative 222;  Mismatches 500;  Indels 172;  Gaps 32;

QY  7 KRPNDGSESS-----SHPPP-----EKETEGSISDYLRIPIRYADKYDNLNIALICA 56
DB 10 KLPKSPGSGTTSCHSVSHAEVLDRLQHLTPVSGIOPGPIYRYATRDVAALLGSLAA 69
QY  57 IGSASLPLMSITIGSFNKNFNNGSGSPFAKADVDHFLVFLFVFLFVFLFVFLFVFLV 116
DB 70 IAGGAALPLFVFLFGLTSTFODIATHRTTYDFHFLHETKNNVFIYILGAAEFVAYLAT 129
QY 117 AAITISAITRTFLRRVELECTLRQEVNHFEDKOSGAIAATXVTTNGNRIQTGIAEKLVFT 176
DB 130 VGIYTGHDHVQOIIRVEYFQALIKONTAFEDTLGAGEITRTIADTNLIQDISEKYGLA 189
QY 177 VQALSMPFSAFVALASOWKALITMSVIPAFVLTGICIAIDAAQEARITRIYSRAVL 236
DB 190 LTGLSTFTVTAFTIAYIKWKLALICSASLLALLTWTGCGSTLMUJFSKKALEYOGGASM 249
QY 237 AEEVLSSIRTHAFYAAQKWKYDVPFLOQAHOEGKKSPNNGVLFSTFEPFCIYAATA 296
DB 250 AEDILDSIRTVAFNAQETLARKYESHLKDAEGPGMKSKVIFAIMVGALLCIMYINGIG 309
QY 297 FNGKFMYONGEVADV--GKVFVTVLTVLAATISLAPSGSVVYQRRIFGSELSFIID 354
DB 310 FNGSFLVEG--ISNIKAGDVLTIMAILLGSYNGNVPNGQALSDAFAAASKLYGTID 368
QY 355 KPTQLDPLDPGKOPGCGLOIEIONLAFAYPSRPSAQVLRDNLTIPAGKTTALVGASG 414
DB 369 RQSPDLALSDQGTLEFFYRGVNLQNIHRHYPSRPEVTVAHDLSCYIPAGKTTAFVPGSG 428
QY 415 SGKSTMGVLEWYLPSSGRILLDGLGOYNVKNLRSIRLVQOEPVLFPGTTFQNIAN 474
DB 429 SGKSTISLIRFYDPVAGTILMDGHDIQTLNLRLWLRQOQSLVSQEPRLFATTIENIRY 488
QY 475 GPMDEORDLPREKQEL---VQAKASGDVFINELPNQSYETVEGERAGALSGGORI 531
DB 489 GIISGR--PEKESYIEIKRVEAARNAHDFIMALPNGYDNI--ESFSLSGGQKQRI 544
QY 532 AIARSIIDPKILLDEATSDALPKAEVVOEALNRVSDRTTLVIAHKLATVKSAGNIA 591
DB 545 AIARAIIDPKILLDEATSDALPKAEVVOEALNRVSDRTTLVIAHKLATVKSAGNIA 604
QY 592 VISQKIVBQGTTHHELIEFGCHYAAALVRAQDLGADQEQHEKTLHEKAAREAAAGERPALE 651
DB 605 VLANGOIVBQGPHEHLNDRRGYCDMVEAHEI-----KKRYSRYSKRYSQSLTNLS 655
QY 652 RTHVTAT-----SQAGDL-----BKRKVPVPTGLYS-LLKCLILIM 685
DB 656 PRHNPMTFFDKYDPDBDESIIYSILSDASDGLHTGEKQ--PVSRMSLSHLAQPVKEE 714
QY 686 FYEQKNLY-----WCLLSTIIVLICAAITF--PGQALLFSRLLVPE---LSGHAA 731
DB 715 AYSFTWTLFELASFRNRPWPELGLUGCASILAGIOPSOAYLFAKAVSTLSLPLEYPKL 774
QY 732 QERADYILMFTFVALGNLVGYFTIGWTCNVISQVTHRYQAAMFORVLDODIELLOIPE 791
DB 775 RHDANFWCLMFLMIGIVSLVLYSVQGTFLFAYSSERKMYARSQAQFVTLHDQIDFFPDQOE 834
QY 792 QISGALTSOLSALPTQLOE-----LISANFL-----IYIVVGQH-----RL 827
DB 835 NTTGALTATLSAGTKELTGISGVTLGTILVSVNLVASLGVALLGWKLALVCISAVPAL 894
QY 828 EDCSTSLWNETGPGGCWCTSTPAFGNLQPNSSRDEARSRLGLCKRCKWACK--RSYR 886
DB 895 LACGFVRVM-----LRFQRPKAKAYQESASS--ACCAASAIR 931
QY 887 DDPRLIFDS-----RRPCSPVTLGHVEQGLAKIITQSPFWGRCGFHLSQSMFEFLA 936
```

```
DB 932 TVVSLTMEATEALQYQAOLRRQLKSDILPIVKSLL-----YASSQALPFFC 978
QY 937 IALGFCAIYDNNWL-----OVSTTQLNFNFISSWAPCLPVOAAAQYLAYSTSFTRKASA 988
DB 979 MALGF-----WYGSLLGHGEYSLFQFYCFSEVIP--GAQAAGTVFESHAPDMGKAKHA 1030
QY 989 ANYILMLRTLKKTIRE--TENKKKG--PVGCGP--VDLEIDIEFRYKQDRARVLGVSMST 1043
DB 1031 A-----REFKLFSSDTMHASRSKGPVVTSMRGLVEFRDVSFRYPSRLEQPLRHNLMT 1084
QY 1044 IEPQFQVAVYVAGSGCGKSTLLJALSERFVDPDTSGRISFAHENIAEMSPRLYRGHMSLVQOE 1103
DB 1085 IKPQFQVAVYVAGSGCGKSTTALLERYDPLKGGVYVDGKNITILEMSSYRSHLALISOE 1144
QY 1104 PTLYQSVRENTVLALEAE--LSPELQCGRLPAPMLNILLSLYQKALKRAAQGHQPS 1161
DB 1145 PTLFOGTIRENILLGNSNTHVTDPLVKACKDANIYDFILS--LPQGFVTIVGNKGML 1202
QY 1162 GGQRQRIATARALIRNPKLLILLDEATSDALQTSERLVQALDEASTRTTIAVAHRLSTI 1221
DB 1203 GGQRQRIATARALIRNPKLLILLDEATSDALQTSERLVQALDEASTRTTIAVAHRLSTI 1262
QY 1222 RNVQVIFVFPANGRIETGTHAEQLRGRVYVCMCAQSLD 1261
DB 1263 QRADLIIVLDQGEVVEGSGTHRELLRKKGRYVELVHLQNP 1302
```

```
RESULT 15
US-09-672-810-4
; Sequence 4, Application US/09672810
; GENERAL INFORMATION:
; APPLICANT: STOCKER, PENNY J.
; APPLICANT: STEIMEL-CRESPI, DOROTHY T.
; APPLICANT: CRESPI, CHARLES L.
; TITLE OF INVENTION: P-GLYCOPROTEINS AND USES THEREOF
; FILE REFERENCE: G0307/7018
; CURRENT APPLICATION NUMBER: US/09/672.810
; PRIOR FILING DATE: 2000-09-28
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1283
; TYPE: PRT
; ORGANISM: Macaca fascicularis
US-09-672-810-4
```

```
Query Match      25.7%  Score 1656;  DB 20;  Length 1283;
Best Local Similarity 32.6%  Pred. No. 5.3e-149;
Matches 428;  Conservative 252;  Mismatches 495;  Indels 138;  Gaps 28;

QY  24 EKETECSIDYLRIPRYADKYDNLNIALICALGSCASLPLMSIIFGFTNK----- 76
DB 27 KKERKPTVSVF--SMFYSNWLIDKLYMVVGTUAAIHGAGLPLMLVFGDMTDTFANAGNL 85
QY  77 -----FNNYNSG---DGSPEA--FKADVDHFLVFLFVFLFVFLFVFLFVFLV 125
DB 86 GDGLALLFNNTNNSNITDTPVMNLEEDMTRYAYIYSGIGAGVLVAAYIQVSFWCLAAGR 145
QY 126 TTTLRARVLECTLRQEVNHFEDKOSGAIAATXVTTNGNRIQTGIAEKLVFTVQALSMFSS 185
DB 146 QIHKIRKOFFHAIMRQEIWDVHVGELNTRLTDDVSKINEGIGDKIGMFFQSMAFET 205
QY 186 AFVVALASOWKALITMSVIPAFVLTGICIAIDAAQEARITRIYSRAVLAEEVLSIR 245
DB 206 GFIVGTRGKWLTVILAIISPVLGSAVWAKILSSFTDKELLATKAGAAVEEVLAIR 265
QY 246 TVHAFYAQKWKYDVPFLOQAHOEGKKSPNNGVLFSTFEPFCIYAATAALAFWKGFRMYQ 305
```

Search completed: April 27, 2002, 07:55:43  
Job time: 604 sec

DB 266 TIIAFGGOKKELERYKNNLEAKRIGIKKAITANISIGNAFLLIYASVALAFWGTTLVL 325  
QY 306 NCEVADYKVFVTVLSTLAATSISMLAPSGSVYQRRIFGSELESIIDRPTQLDPLDPS 365  
DB 326 SKEYS-IGQVUTVFFSVLIGAFVGOASPSIEAFANAGAFAFEIKIIDNKPSIDSYSKS 384  
QY 366 GKQPEGCLIGETONLAFAYPSRPSAQVURDFNLTPAGKTTALVGASGCKSTMVGLLE 425  
DB 385 GHPDNKGNLEFRNVHFSYPSRKEYKILKGLNLKVSQGVVALVGNCGCKSTVOLMQ 444  
QY 426 RWYLPSSGRILLDGLGQVNVKWLRSRLVQOPVLFGRGTIFONIANGFMDQRDLPR 485  
DB 445 RLYDPTGMSVVDQODIRTINVRFLIIGVVSQEPVLFATTIAENIRYGRD----- 497  
QY 486 EKOMELVQKACKASNGDVFNEPNCYFTEVGERAGALSGGORORAIARSIIISDPKILL 545  
DB 498 -VTWDETEKAVKANAYDFIMKLPQKDFLVGERGAQLSGGOKOKIAIARALVRNPKILL 556  
QY 546 LDEATSAIDPKAKVVOEALNRVSKORTTLVIAHKLATVKSAGNIAVISQCKIVEQGTTH 605  
DB 557 LDEATSAIDPESAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDGVIYKGNHD 616  
QY 606 ELIEFGCHYALVRAQDLGADEQHEKTLHEKAAREAAGERPALE-RTHTTATSAQOL 664  
DB 617 ELMKEGIFYKLVMTAGNEIELEN-----AADSKSETDLEMSHDSGSSLIRKR 669  
QY 665 EKRVPVVTGLG-----YSLKLCILIMFYEQKNLYW-CFLLSTIV 703  
DB 670 STRSRVSGOQDKLSTKEALDESIPPVSPWRINKNL-----TENPYFVGVFCA 721  
QY 704 LICAATPGGALLFSRLTVELSGHAAQER--ADFYILMEFVVALGNLYCYFTIGTCN 761  
DB 722 IINGLOPAPAVIFSIIIGIFRNDDAETKRONSLFSLFLVLGIVSPTFFLQGFPG 781  
QY 762 VISQVTHRVQAMFORVLDIDPEQISGALTQSLSALPTQJLQELISANFLIYV 821  
DB 782 KAGEILTKRLRYMVFRSMURQVSWFDDPKNTGTALTRLANDAQAQVKAIGSLAI--- 838  
QY 822 VGOHREQCSTSLMMETPGGCVMTCTSPAEGW-----LPQ 858  
DB 839 -----ITONIANLGTGII---SLIYGWQLTLLLAIVPIIAIAGVVMKMLSG 884  
QY 859 NSSREARSRKLGKCKKCKWACKRSYRDPDLIFDSRRPCSPVVLGHV-EOGLA----- 912  
DB 885 QALDKKKELEGACKIATEA---IENFRVVSITQOK-----FEMYDOSLOVPYRN 933  
QY 913 KIQSFVWFCRCFPHLSQSMFLAIALGFCIAVDNWLQVSTQLNFTSSSWAFCLPV--- 969  
DB 934 SLRAHIFGTFSF--TOAMWFSYA--CCFRFGAIL-VAHSLMSFEDVLVFSVAVFGA 988  
QY 970 QAAAQYLAYSTSTFKARSAANYILMLRTLKPTIRE-TEENKKGVPVGGCPVDLEDIEF 1028  
DB 989 MAVGVSVFAPDYAKAKVSAAHIIIMTILDSYSTYGLKPKNTLES-NVTNEVVVFN 1047  
QY 1029 RQRDSARVLKRVSMTEPQGVAVYGASGCKSTLIALSERFYDPTSGRISFAHENIAEM 1088  
DB 1048 PTRLIDIPVLOGLSLEVKKGQTLALVSGSGCKSTVQVLLERFYDPLAGKVLDDGKEIK 1107  
QY 1089 SPRLYRGHMSLVQOEPFLYQSVRVNVTLEAEEL--SEELCOGRLPARPMWILSLYQ 1146  
DB 1108 NVQWLRHLGIVSQEPILFDCSISENIAYGDNRSRVVQVEIVRAAKEANITHAFI-ESLPN 1166  
QY 1147 KALFRLAARQMFGSGGORORAIARALIRNPKILLIDDEATSAIDTQSERLVQALDEAS 1206  
DB 1167 KYSTRV-GDKGTQSGQKQRIARALVRQPHILLDEATSAIDTQSERLVQVQALDKAR 1225  
QY 1207 TSRTTIAVHRLSTIRNVDFVVFANGRIAETGTGHAELORLGRYEMCLAQS 1259  
DB 1226 EGRTCIVIAHRLSTIONADLIVVFQNGRVKHEGTHQQLAQKGIYFSMVSYQA 1278

-----





```
Qy 94 VDH-----FVLMFVILFGKFLVTVYVSTAATISAIKTRTRTLLRRVFELECTLR 140
Db 123 TESQRTQVTPMPSPWRILYFVGLIARLVSYNTLLTYAARIVRNIRHAYLKAAALS 182
Qy 141 QEVWHFKQSCATATVYTNNGRNIQFGIAEKLVFTVQALSMFSAFVVALASQWKALAI 200
Db 183 QEVAYIDFGSGSIAAQATNGKLIQAGASDKIGLLFQGLAAFFVTLSSLRLCKWKLTLI 242
Qy 201 TMSVIPALFVTCICIAIDAOEARITRYSRAAVLAEEVLSSRTVTHAFYAOKKWKVEY 260
Db 243 CICIPVATIGTGVAAVEAGHETRIQIHAQANSFAEGILAGVKAHVAFGMRDLSLRKF 302
Qy 261 DVFLQAQHQEGKKSPNGGVLFSTEYFCIYAALAFKWKGRMYQNGEADVGVKVFVVVL 320
Db 303 DEYLVEARKVKKISPLGLLSAEYTIIVLYGLAPMOGIMHMFGRGEICTAGDIPTVL 362
Qy 321 SVTLAATISIMLAPSGSVYQRRITFGSELFSIIDKPTQLDPLDPGSKQPEGGLQGEIQN 380
Db 363 SVTIASINTLLAPSYSTEFRAASAAQPLRIDRESEINPYKGELEPERVILGDVELEN 422
Qy 381 LAFAPRSPSAQVLDNFNLIPAGKTTALVAGSGSGKSTMVGLLRLWYLPSSGRILLDGL 440
Db 423 VTFSPTRPGITVLDNLSKVPAGKVTALVQSGSGKSTIVGLLRLWYLPSSGAIRLDGN 482
Qy 441 ELGOYNNKWLBSRLVQOEPVLFPGTIFONIANGFMDQ- RDLPRKQHELKQKACKAS 499
Db 483 LISLNVGLWRNRLVQOEPVLFQSGVFDNIRYGLVGTPTWENASREOEQERVQEAALA 542
Qy 500 NGDVFINELPNGYETVEGERAGALSGGQORAIARIISDPKILLDEATSSALDPKAEK 559
Db 543 YAHEFISELTDGDTLIGERGLSGGQKQKQVARIASVVSQPKVLLDDEATSSALDPKAE 602
Qy 560 VQOALNRVSKDRITVIAHKLATVKSAGNIATVISOQKIVQEGTHHELFEGGCHYAALVR 619
Db 603 IVQALDKAABGRITVIAHKLATIRKADNIVNMSKGHVQEGTHESLIAKGVAGLVK 662
Qy 620 AODLGAEQOQHEKTLHEKAAREAGERP- LERTHTTAT-----SQAGDLE 665
Db 663 IONLAVN-----ASAHNVNEEGEDVALLEVTAVTRYPTSIRGRMNSIKDRDYE 716
Qy 666 KRKVPVGTGLYSLKILIMFYEQKNLYWCFLLSTINVLI CAATFPCQALLFSRLITVFE 725
Db 717 NHR-----HMDMLAALAYLVRECPKWAYLVLLCGLCGCAMPQQAILMSRVVEFT 770
Qy 726 LSGHAAGERADFYILMFVVALGNLYGFTIGWTCNVISOVTHRYQAAAFQVLDQDIE 785
Db 771 LSGDAMLDKGDFYASMLIVLAAGGLICYLAVGYATNTIAOHLSHWFRLLIHDMLRDIQ 830
Qy 786 LLDIPEQISGALTSQLSALPTQLOELISANFLIVVQHRLEQCSTTSLMMETGPGCV 845
Db 831 FFDRENTTGALVSRIDSPHAILLEMGYNIAL-VVIAVLQVTCGTLAIFSKWGLV 889
Qy 846 WCTSTPAFGWLPQ-----NSRDRSARSKLG-----KLCRKC 877
Db 890 -----VFGGIPPLVAGMVRIRVDSRLDRQTSKKYGTSSSIASEAVNAIRTVSSLAIEE 943
Qy 878 WACKRSYRDPDLRIFDSRSPCTVLGHVEQGLAKIIQSFMTGRCFCFGLHSQSMFLAI 937
Db 944 TVLRYRT-EELDHAVSSSVKPMATMI-----CFG--LTQIEYWFQ 982
Qy 938 ALGFCIAVDNWLQVSTQNLPISSWAFCLPVQAAQYLAYSTSFTKARSAANYILWRT 997
Db 983 ALGFWYCGRLVSLGETSNYSFFVAFLSVFFAGQASQALFQWSTSTKGINATVIAWLHQ 1042
Qy 998 LKPTIRTEENKKGPGVGGCPVLEDEYEFYRORDSARVLRGVSWTTPQGFVAYVGASG 1057
Db 1043 LQPTVRETPENHDKPGSGAPIANDNVFSPYLRPDAPILKGNVLMKINKGOFINFGSSG 1102
Qy 1058 CGKSTLALSERFDPSTSGRISFAHENIAEMSPRLYHGMHSLVQOEPFLYQGSVRENTL 1117
Db 1103 CGKSTMIAMLERFDPTGTSTIDASTLTDINPTISYRNIVALVQOEPFLYQGSVRENTL 1162
Qy 1118 A-LEAEELSEELQGRLEPAPMLWTLSSLYQALKRLAAQGMQFSGGQORAIARALIR 1176
```

```
Db 1163 GDAVKSVDQIESALRAANAMDFVSSLPQ-GIYTPAGSGGSQLGGQRIARALIR 1221
Qy 1177 NPKLLLDDEATSSALDQTSERLVQAAALDEAST--SRITIAVAHRLSTIRNVDFVFPANGR 1234
Db 1222 DPKILLDDEATSSALDQTSERLVQAAALDEAST--SRITIAVAHRLSTIRNVDFVFPANGR 1281
Qy 1235 IAEFTGHAELQRLGRYRYEMCLAQSLD 1261
Db 1282 IAEFTGHAELQRLGRYRYEMCLAQSLD 1308

RESULT 2
US-09-352-552-2
: Sequence 2, Application US/09352552
: Patent No. 6060264
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Waard, Maarten
: TITLE OF INVENTION: Multiple Drug Resistance Gene atnC of
: TITLE OF INVENTION: Aspergillus nidulans
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09352.552
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/996.644
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11765
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1308 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-352-552-2
```

Query Match 35.2%; Score 2270; DB 3; Length 1308;  
Best Local Similarity 39.8%; Pred. No. 1.8e-228;

Matches 512; Conservative 225; Mismatches 438; Indels 112; Gaps 19;

```
Qy 46 WTLNVTALICA-----IGSGASL-PLMSIIFGSFTN--KFNYNNSGDSGSPFAKAD 93
Db 63 WRIQAKLTCRSKRGSLHSPAGONSFRPLLSLLHAPLEQELRPFKTSASSSPSPISP 122
Qy 94 VDH-----FVLMFVILFGKFLVTVYVSTAATISAIKTRTRTLLRRVFELECTLR 140
Db 123 TESQRTQVTPMPSPWRILYFVGLIARLVSYNTLLTYAARIVRNIRHAYLKAAALS 182
Qy 141 QEVWHFKQSCATATVYTNNGRNIQFGIAEKLVFTVQALSMFSAFVVALASQWKALAI 200
Db 183 QEVAYIDFGSGSIAAQATNGKLIQAGASDKIGLLFQGLAAFFVTLSSLRLCKWKLTLI 242
```



QY 201 TMSVPAIPFLVTGICIAIDAAOEARITHIYSRAAVLAEEVLSIRTVHAFYAKKWKVY 260  
Db 243 CICIPVATIGTGVVAAVEAGHETRIQIHAQANSFAEGLAGKAVHAFGMRDLSLRKP 302  
QY 261 DVFLOQAHOEGKKSPNGVLFSTFYTCIYAAIALAFWKGRFYONGEVADVGVFTVYL 320  
Db 303 DEYLVFAHVKVGGKISPLGLLFSAEYTIYILGYLAFWQGIHMFGRGEICTAGDITFVLL 362  
QY 321 SVTLAATSISMLAPSGSVVYORRIFGSELSFIDKPTQLDLPDSKOPGCGIGLEION 380  
Db 363 SVVIASINLTLLAPSYIESFSAASAAQLFRLIDRESEINPYGKEGLEPERVCLDVELEN 422  
QY 381 LAFVPSRPAOVLKDFNLITPAKTTALYVAGSGSGSTWYGLERWYLPSSGRILLDGL 440  
Db 423 VTFSTYPTRGITVFNLSKVPACKVTLVQSGSGSTVGLERWYNTSGAIRDGN 482  
QY 441 ELGQYVNWLRISRLVQOEPVLRFTGTFQIANGFDEQ-RDLPREKQWELVQKACKAS 499  
Db 483 LISLNVGLNRNRLVQOEPVLFQGSVFDNIRYGLVGTWENASREQMERVOEAAKLA 542  
QY 500 NGDVFNELPNGYETEVGERAGALSGGORBIATARSIIIDPKILLDEATSLDPAEK 559  
Db 543 YAHEFISELTDGTDILGERGLSGGQKQVARSVQPKVLLDEATSLDPAEAT 602  
QY 560 VVOEALNRYSKORTTLVIAHLKATVKSAGNTAVISOGKIYVQEGTHHELIEFGCHYAALVR 619  
Db 603 IVQKALDKAAGRTTIVIAHLKATIRKADNIVMSKGIHVEQGTSHESLIJAKDGVYAGLVK 662  
QY 620 AODLGADEQOEHEKTLHKAAREAGERPA-LERTHTTAT-----SQAGDLE 665  
Db 663 IONLAVN-----ASAHNVNEEGEDEVALLEVTETAVTRYPTISIRGRMNSIKRDQVE 716  
QY 666 KRKVPVGTLYGSLKCLILMFYEKNLYWCFLSTTVILCAATFPQOALLFRLTVFE 725  
Db 717 NHR-----HMDLAALAYLVRECELEKWAYLVLLCLGSCAMYPQOATILMSRVVEVT 770  
QY 726 LSGHAAQERADRYIIMFFVVALGNLVGYFTIGTCNVIQSVQVTHRYQAAAFORVLQDDIE 785  
Db 771 LSGDMLDKGDFVASMILVLAAGCLICYLAVGYATNTIAQHLSHWRRLILHMDLRQDIQ 830  
QY 786 LLDIPEQISGALSQLSALPTOLQELISANFLIVVQHLQHEQCSTSLWMETGGCGV 845  
Db 831 FFOREENTGATVSRIDSPHAILDEMGYNAL-VVIAVLQVVTGGLATAFWSKGLGVV 889  
QY 846 WCTSTPAFQWLPO-----NSSRDEARSRLG-----KLRKC 877  
Db 890 -----VFGGIPPLVAGNVRIRVDSRLDRQTSKKYGTSSIASEAVNAIRTVSSLAIE 943  
QY 878 WACKRSYRDPRLIFDRRPCSPVLGHVBOGLAKIIQSPWGRFCFGFHLSDSMFLAI 937  
Db 944 TVLRYT-BELDHAYSSVKPNAATMI-----CFG--LTQCIEYWFQ 982  
QY 938 ALGFCIANDWLVQVSTTQLNFTISSWAFCLPVQAAQYLAYSTFTKARSAAANYILWRT 997  
Db 983 ALGFVYGCRLVSLGTSMTYSFVAFLSVFFAQASQALFQWSTITKGINATYIAWLHQ 1042  
QY 998 LKPTIRETENKKGPVGCGPVDLEDIEFRYRQDSARVLGVSMTTEPGQFVAYVAGSG 1057  
Db 1043 LQPTVRETENHDKPGSGAPITAMDNRFSYPLRPDAPILKGVNLKINKGOFIAFVSGSG 1102  
QY 1058 CGKSTIALSERPYDPTSGRISFAHENTAEMSPLRYCHMSLVQOEPFLVQGSVRENTL 1117  
Db 1103 CGASTMIAMLERPYDPTSGITSIDASTUDINPSYRNIVALVQOEPFLVQGTIRDNISL 1162  
QY 1118 A-LEAESEELCOGLPARPMLWILSSLYQKALKRLAQRGMQFSGGORIARIALR 1176  
Db 1163 GDAVKSVSDEQIESALARANAWDFVSSLPQ-GIYTPAGSGGSLGSGGORIARIALR 1221  
QY 1177 NPKILLDEATSLDPTQSERLVOAALDEAST--SRTTIAVAHRLSTIRNVDVIFVPANGR 1234  
Db 1222 DPKILLDEATSLDPTQSEKIVQKALEGAARDGDELIVVAHRLSTIKDANVICVFGGK 1281  
QY 1235 IAEFTGHAEQLRLRGRYEEMCLAQSLD 1261

Db 1282 IAEFTGHAEQLRLRGRYEEMCLAQSLD 1308  
RESULT 3  
US-08-612-734B-2  
; Sequence 2, Application US/08612734B  
; Patent No. 5914246  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: Tobin, Matthew B.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene of  
; TITLE OF INVENTION: Aspergillus Fumigatus  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center, DC1501  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.A.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,734B  
; FILING DATE: 08-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Craig, Anne I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: X-9681  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-354-4043  
; TELEFAX: 617-354-9570  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1349 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-612-734B-2

Query Match 30.3%; Score 1956; DB 2; Length 1349;

Best Local Similarity 35.8%; Pred. No. 2e-195;

Matches 491; Conservative 229; Mismatches 478; Indels 172; Gaps 29;

QY 4 ESEKPRNQDGSSESSHPPP-EKETEGSISD----- 33

Db 35 DNEAPHDHLSUDTIMAPPDCKRDKGKAVDLDNDSLFAHLQHEKEVLRKOLDAPSVKV 94

QY 34 -YLIRFYAOKYDWTNLVIALICAGSCASLPLMSIFGFTNKNPNYNSGDSPEAFKA 92

Db 95 SFTLYRYASRKDILLIILVSAICAIAAGAAPLFTILFGSLASAFQIGSLGTMPTHEFYH 154

QY 93 DVDFHVLWVYLFITGKVLVTVVSTAAITTSIAITRTTLRRVFLCTRLQEVWHFQKQSG 152

Db 155 KLRNVLYFVYLGTAERTVTVSVTVGVPIYGEHLTKIRENYLEAILRONMAYFDKLGAG 214

QY 153 ALATXVTNGNRIQTGAELVTVQALSMFFSAFVVALASQWKLALITWVIPALFLVT 212

Db 215 EVYTRITADTNLIQDAISEKVGTLTAFATVFAIVAYVYKWKALICTTIVALVVMV 274

QY 213 GICIAIDAAQEARITRIYSRAAVLAEEVLSIRTVHAFYAKKWKVYDFVFLQQAHOEKG 272

Db 275 GGGSRFIVKYKSKSIESYAGGTVAEERVISSIRNATAFGTQDLAKOYETHLAEAEKMGV 334

QY 273 KSPNNVGLFSTFECFIYAAIALAFWKGFYQNGEVADVGVFTVTVLSTLAATSIHNL 332

Db 335 KQVILGMLGGNFGTMSNYGLGFWMGSRFVYCKEV-NVQCVLTVLMSILIGSFSLSGNV 393  
QY 333 APGSGVYVYQRIIPGSELSFIIDKPOLDPLDPSGKQPGCGIQTGIONLAFAYPSRPSAQ 392  
Db 394 APNGQAFITGVAAGAAKIYTIIDRRPLDPSDEGKVLDFHFCNTEFRNVKHIYPSRPEVT 453  
QY 393 VLDFNLITPAGKTTALVAGSGSGKSTWVGLERWYLPSSGRILLDGLGQYNVWKLRS 452  
Db 454 VMEDVSLSPAGKTTALVAGSGSGKSTVVGVERFYLPGVGVLDDHDIQTILNLRWLRQ 513  
QY 453 RIRLVQEPVLEGTETQNIANG-----FMDEORDLPREKQOMELVOKACASNGDVFINE 507  
Db 514 QISLVQEPVLEGTETQNIANG-----FMDEORDLPREKQOMELVOKACASNGDVFINE 507  
QY 508 LPNGYETEVGERAGALSQQORQRIAIARSIIISDPKILLDPAISALDPKAEKVVOEALNR 567  
Db 570 LPEGYDTWVGQFLLSGGKQRIAIARIVSDPKILLDEATSAIDTKSEGVQOALDK 629  
QY 568 VSKDRITLVIAHKLATVKSAGNIATVISOQKIVQEQTHLIEFGCHVAALVRAQ----- 621  
Db 630 AASGRTTIVIAHRLSTIKTAHNVAMVGKTAEOGTHDELVDKGTYYKLVQAEQRINEEK 689  
QY 622 -----DLGADEQOEHEKTLHEKA-----AREACERPALETHIT-TATSOAGDLEK 666  
Db 690 EABALEADADMDADFGQGVTRITATVSSNSLDAVDEKARLEMKRTGTOKSVSSAVLS 749  
QY 667 RKVPVGTGLYSLLKCI-LIMFYEQNLWYCFLLSTITVLIACAATFPQALIFSRLLTVE 725  
Db 750 KKVPEQEKYSVTLVKFTAGFNRPELGYMLIGLTFSP-L-AGGQOPTQAFYAKAISTLS 808  
QY 726 LSG---HAAQERADFIILFVVVVVGLVGNLVGVTIGTWCNVISQVYVTHRYQAAAFORVLDQ 782  
Db 809 LPESMFHKLHNDANFSLMFFVVGTAQFISLSINGTAFACSEBRLIRARSQAFRSILRQ 868  
QY 783 DIELLDIPGISCALTSQISALPTQLOEL-----ISANFLYIVVVGQRLE 828  
Db 869 DISFFDRENSTGALTSFSLSTETKNSLGSVGVLTGIIIMTSTTLGAAMITALAGKLLAL 928  
QY 929 QC-STTSLMNETGPGCVCTSTPAFGWLPQNSSRD-----EARS--RKLGLK 873  
Db 929 VCISVVPILLACG-----FLRFYMLAQFOQRKSAYEGSASAYACEATSARTVASL 979  
QY 874 CRK--CWACKRSRYRDPDLFDSRRPCSPVTLGHVEOGLAKIITOSFWFGRCFGHLSQS 931  
Db 980 TREQWGVVYHDQLOQGR-----KSLISVLRSSL-----YASSQA 1016  
QY 932 MEFLAIALGFCIAVDMWLOVSTQLN-----FSSSWAFCLPVQAAAQIAYLSTSF 982  
Db 1017 LVFFCVAGLP-----W--YGGTLLGHHEYSIFREFFVCF-S-EILFGAQSAGTVFSFAPDM 1067  
QY 983 TKARSANYILWLTLPKPTIR-ETBENKKGPVGGCPVDLEDIBFPRYQRDSARVLRGVS 1041  
Db 1068 GKANMAAQPKLFDKSKPTIDWSEGEKLESMEG-EIEFRDVHFVYRTPRPEQPVLRGLN 1126  
QY 1042 MTPEGOFVAYVASCCKSTIALSERPYDPTSGRISFAHENTAEMSPRLYCHMSLVQ 1101  
Db 1127 LSVAGKRLTALVGCCKSTIALSERPYDPTSGRISFAHENTAEMSPRLYCHMSLVQ 1186  
QY 1102 QEPFLYOGSVRENTVLALEA-ELSEE---LQQRHLPARPMWLISLIYQ-----KALK 1150  
Db 1187 QEPFLYOGKILNLLGVKDDVSEETLIKVKCD-----ANLYDFVMSLPEGFD 1235  
QY 1151 RLAQRCMQFSGGQRIAIARALRNPKLLLDDEATSAIDTQSERLVQOALDEASRT 1210  
Db 1236 TVVCSKGLMSGGKQORVALARALLRDPKVLLEDATSAIDTQSERLVQOALDEASRT 1295  
QY 1211 TIAVAHRLSTIRNVDFVFPANGRTAEVGTHTAELQRLGRYEMKLAOSL 1260  
Db 1296 TIAVAHRLSTIQADIIYVFDQKIVSGSTHRELIRNKRGRYELVNLQSL 1345

RESULT 4  
US-08-996-545-2

; Sequence 2, Application US/08996545  
; Patent No. 5928898  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
; TITLE OF INVENTION: Aspergillus Nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,545  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1334 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-996-545-2

Query Match 30.1%; Score 1940; DB 2; Length 1334;  
Best Local Similarity 36.8%; Pred. No. 9.5e-194;  
Matches 490; Conservative 231; Mismatches 483; Indels 128; Gaps 30;

QY 2 ADESEKPRPDGSGESS-----SHPPPE-----KETEGISDYLRIFRYADKY 44  
Db 54 ADKEHPRKSSSSNNVSVNEVDALIAHLPEDEHQLKTOLEEKVNIS-FFGLMRYATKM 112  
QY 45 DWTLNVALICAGSASLPLMSIIFGSFTNKNYNSGDGSPKADVDHFLVLFYVL 104  
Db 113 DILIMWISTICATAAASFTQRIIMLYQISY-----DBFYDELTKNLYFYVL 158  
QY 105 FIGFVLTYVSTAATITSAITRTTLRRVFLCTLRQEVYHFDKOSGAIAFXTVTNGNR 164  
Db 159 GIGFVTVYVSTVGTPIVYGEHATQKIREYLESILRONICYFDKLCAGEVTRITADTNL 218  
QY 165 IOTGIKALVPTVOALSNFSAFVALASQWALITMSVIPAIPLVTCICIAIDAAQEA 224  
Db 219 IQDGISEKVGTLTALATFTVTAIIYVYVKLALICSSSTIVALVLTWGGGSGQFIKYSK 278  
QY 225 RITTIYRAAVLABEVLSSIRTVAFAQKVKVEYDVFLOQAQHECKKSPNNGVLFST 284  
Db 279 KSLDSYGAGGTVAEEVLISSIRNATFCTQDKLAKQVEVHDEAKMGTKNQIVMGFMIGA 338  
QY 285 EYFCIYAAIALAFKGRMYQNGEVADGVKFTVVLSTLAATSIKSLAPSGSVVYQRI 344  
Db 339 MFLMYSNYGLGFWMGSRFLVDGAV-DVGDILTLYMAILIGSFLGNSVSPNAQAFTNAVA 397  
QY 345 FGSELSFIDKPTOLDPLDPSGKQPGCGIQTGIONLAFAYPSRPSAQVLRDNFTIPAG 404  
Db 398 AAAKIFGTIDRQSLDPSYSGEKTLOHFEGHIELRVKHIIYPSRPEVTYMEDVSLSPAG 457

QY 405 KTTALVGASGSGKSTWVGLLRLRYWLVSSGRILLDGLLELCOYVNVKWLRSRIRLVQOEPVLF 464  
Db 456 KTTALVGSGSGKSTWVGLVREYFVMPVGTVLVDHDKDLNLRQOISLVSOEPVLF 517  
QY 465 RGTIFONTAGFMDEQDRLPRE-KOMELVQKACKASNGDVFINELPNNGYETEYVGERAGAL 523  
Db 518 GTTIIKNIRHGLIGTKYENESDKVRELITENAKMANAHDFITALPEGYETNVGQGFLL 577  
QY 524 SGGORORIATARSITISDPKILLDEATSALDPKAEKVQOELNVRVSKORTTILVAHKLAT 583  
Db 578 SGGQKORIAARAVSDPKILLDEATSALDPKSEGVQOALERAEGRTIIVIAHRLST 637  
QY 584 VKSAGNIATVISOQKIVEQGHHELIIEFGCHYAALVRAODLGADEQOEH- 632  
Db 638 IKTAHNVLVNGKIAEOGHDELVDVGGAYKRLVQAQRI--NEQKEADALFADAEHLT 695  
QY 633 -----KTLHEKAAREAAAGERPALERTHTATSOAGDLEKRVPGVGTGLYSLLKCI-LI 684  
Db 696 NADIAKIKTA-SSASSDLGKPTTIDRTGTHKSVSAILSKRP-PETTPKYSLWTLKLFV 753  
QY 685 MFYBOKNLYWCFLISTTIVLCAATFPGQALLFSRLTYFEL---SCHAAQERADFYILM 741  
Db 754 ASFNRPETPY-MLGLVFSVLAGGQPTQAVLYAKAISTLSLPSQYSKLRHDADFWSLM 812  
QY 742 FFVALGNLVGYFTIGWTCNVISQVYTHRYQAMFORVLDQIELDIPQISGALTSOL 801  
Db 813 FFVVGIIOTFQSTNGAAFAVCSEIRIIRARSTAFRTILRQDIAFFDEKNSGTALTSFL 872  
QY 802 SALPTQLOEL-----ISANFLIYIVGOHRELOC-STTSLMETPGGCVW 846  
Db 873 STEFKHLSGVSVGLTILMTSTLGAIIIALAIGWKLALVCISVVPVLAG- 926  
QY 847 CTSTPAGWLPQNSRDEARSRLKGLCRKCHACKRSYRDPDRILFDSRPPCSTVLGH 906  
Db 927 ---FYRMYLAQFOSR-----KLAYEGSANFACEATS---SIRTVASLTRERDWEIYH 975  
QY 907 VE---QG---LAKIQSFWEGRGFGPHLSQSMFEFLAIALGFCIAVDNMLQVSTTQLN--- 957  
Db 976 AOLDAQHTSLISLVRSLL-----YASSQALVFCVGLGF-----W-YGTLGLGHHE 1022  
QY 958 -----FISSSWAPCLVQAAQYLAYSFTFKARSAANYILWRLTKPTIRE-TEENKK 1010  
Db 1023 YDIFRFFVCFPS-EILFAQAGVTFSPADMGKAKNAAEFRRLDRKPDQIDNWESEGEK 1081  
QY 1011 KPGVGGCPVLELIEFRYQORDSARVLRGVSMTIEPGVFVAYVAGSGCGKSTLIALSERF 1070  
Db 1082 LETVEG-EIEPRNVHFRYTRPEQVPLRGDLTVKPGQVVALVPGSGCGKSTLIALSERF 1140  
QY 1071 YDPTSGRISFAHENIAEMSPRLYRGHMSLVQOEPVLYOGSRENTVTLAL-EAELSEE-LC 1128  
Db 1141 YDAGSLVLDGDKISKLINSYRSLVSGQPTLYQGTIKENILLGIVEDDVPEEFLI 1200  
QY 1129 OGRUPARMLWILSSLYQKALKRLAQRGMQFSGGQRIATARALRNPKILLDEATS 1188  
Db 1201 KACKDANIYDFIMS--LPEGFNTVWVGSGGMLSGGQKQVJAJARALLRDPKILLDEATS 1258  
QY 1189 ALDTQSERVQAALDEASTRTTIAVAHRLSTRNVVDVIFVEFANGRIARTGPHAELORLR 1248  
Db 1259 ALDSESEKVVQAALDAARCTTIAVAHRLSTIQKADVIYVFDQKIVESGTHSELVQKK 1318  
QY 1249 GRYYEMCLAQSL 1260  
Db 1319 GRYYELVNLQSL 1330

## RESULT 5

US-09-328-320-2  
; Sequence 2, Application us/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.

; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/328.320  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996.545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1334 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-328-320-2

Query Match 30.1%; Score 1940; DB 4: Length 1334;

Best Local Similarity 36.8%; Pred. No. 9.5e-194;

Matches 490; Conservative 231; Mismatches 483; Indels 128; Gaps 30;

QY 2 ADSEKPRNODGSESS-----SHPPPE-----KETECSIDYLRIPRYADKY 44

Db 54 ADKEHRPKSSSSNNVAVSNEVDALIAHLPEDEQRQVLKQLEEKVNIS-FFGLWRYATKM 112

QY 45 DMTLVNIALICAGSGASLPLMSIIFGSFTNKNFNNSGDSGSPFAKADVDHFWLWVYL 104

Db 113 DILWIVISTICALAAATFORIMLVQISY-----DEFYDELTKNVLVYVYL 158

QY 105 FICKFVLTVYSTAAITISAIRTRRLRRVFLCTLRQEVWHFQKSGNGATATXVTTNGNR 164

Db 159 GIGEFVTVVSTVGTIYGEHATQKIRYLESLRQINIGYFDKLGAGEVTRITADTNC 218

QY 165 IOTGIAEKLVFTVOALSMFSPFVVALASOMKALITMSVIPALFVLVTGICIAIDAAQEA 224

Db 219 IQDGISEKVLTLTALATVTAIILAYVKNKALICSTIVALVLTMTMGSGSQFIKYSK 278

QY 225 RITRIYSRAAVLAEEVLSIRTVHAFYAKKMWKEYDVFLQAOHGEKKKSPNNGVLEST 284

Db 279 KSLDSYGAGGTVAEEVSISSIRNATAFGTQDLAKOYEVHLEAEKWTGKNOIVMGPIGA 338

QY 285 EYFCIYAAIALAFWKGFRMYQNGEADVGVFTVSVLSTLAATSISMLAPSGSVVYORRI 344

Db 339 MFGLMYSNTGLGFWMGSRFLVDGAV-DVGDILTIVMAILIGSFLGNVSPNAQFTNAVA 397

QY 345 FGSELSIIDKPTQDLPDPSGQKQEGCLQIEIQNLAFAYPSRPSAQVLRDNLTIPAG 404

Db 398 AAKIEGTIDRQSPDPYPSNECKTLDHFEHIELRNKHIYPSRPEVTMEDVSLSPAG 457

QY 405 KTTALVGASGSGKSTWVGLLRLRYWLVSSGRILLDGLLELCOYVNVKWLRSRIRLVQOEPVLF 464

```

Db 458 KTTALVPGSGSGKSTVGLVVERFMPVGVLLDGHDKDLNLRLQOISLVSEPVL 517
QY 465 RGTIFONTANGFMEQDRLDPE-KOMELVOKACKASGVDFINELPNGVETVGERAGAL 523
Db 518 GTTIYKIRHGLIGTKYENESDKVRELIENAAKMANAHDFITALPEGYTVNGQGFLL 577
QY 524 SGGQRQIARSIISDPKILLDEATSLDPAEKVQOEALNRVSKDRTTTLVIAHKLAT 583
Db 578 SGGQKQRIARAVVDPKILLDEATSLDTPKSEGVOAALERAAGRTTIVIAHRLST 637
QY 584 VKSAGTAVISQGIQSGPHHELIEGCHYAALVPAQDLGADQOEHE-----632
Db 638 IKTANHIVLVNGKIAQGGHDELVDGGAIRKLVAQRI--NEQKEADALDADAEDLT 695
QY 633 -----KTLHEKAAREAAGERPALERTHTTATSOAGDLEKRPVGTILGYSLKCI-LI 684
Db 696 NADIAKIKTA-SSASSDLGKPTTIDRTGTHKSVSSAILSKRP-PETTPKYSLWTLKAV 753
QY 685 MFYEOKNLYWCFLLSTITVICAATFPGQALLPSRLTVFEL---SGHAAQERADFTILM 741
Db 754 ASFNEPEIPY-MLIGLVSVLAGGGQPTQAVLYAKAISTLSLPESQYKLRHDADFWSLM 812
QY 742 FFVALGNLYGYFTIGTCNVISGVYTHRYOAAAMFORVDDQIELDIPQISGALTSOL 801
Db 813 FFVGIQIFQISQNGAAFAVCSERLIRARSTAFRTILRODIAFDKEENSTGALTSLF 872
QY 802 SALPTQLQEL-----ISANFLIYIVVQGHRLQEC-STTSLMMETGPGGCW 846
Db 873 STETKHLSGVGVTLGITMTSTTLGAAIIIALAIGKALVALCISVVPVLLACG-----926
QY 847 CTSTPAFCWLPQNSRDEARSRLGKLCRCWACKESSYRDPORLIFDSRRCPSPTVLGH 906
Db 927 ---PYRFVWLAQFSRS-----KLAYEGSANFACEATS---SIRVVASLTRERDWEIYH 975
QY 907 VE---QG---LAKIIQSFWRGFCGPHLSOSMEFLAIALGFCIAVDNWLQVSTTOLN---957
Db 976 AOLDAQGRTSLIVLSRL-----YASSQALVFFCVLGF-----W-YGGTLLGHHE 1022
QY 958 -----FTSSWAFCLPVQAAOVLAYSTSFTKARSAANVILWLRTLPKPTIRE-TEENKK 1010
Db 1023 YDIFRFFVCFES-EILFCAQASAGTVSFAPDMGKAKNAAEFRRLPDRKPOIDNWESEGEK 1081
QY 1011 KGPVGGCPVLEDEFFRYRORDSARVLGVSMPIEGQVAVVAGSGGCKSLIALSERP 1070
Db 1082 LETVEG-EIEPRNVHFRPTREPQPLRGDLDTLVKQGVVALVGPVGGCKSTIALLERF 1140
QY 1071 YDPTSGRISFAHENIAEMSPRLYRGHMSLVQOEPYLYOGSVRENTVAL-EAELEE-LC 1128
Db 1141 YDAGSTLVQDKDISKLININSYRSFSLVSOEPTLYOGTIXENILLGIVEDDVPPEFLI 1200
QY 1129 OGRLPARPMLWILSSLYOKALKRLAORGMOFSGGQORORATARALIRNPKILLDEATS 1188
Db 1201 KACRDANIYDIMS-LPEGFNTVSGKGLSGGQKQVATARALLRDPKILLDEATS 1258
QY 1189 ALDQOSERLVOALDEASTGRTTIVIAHRLSTRINVDVIVFVANGRIATGTHAEQLRLR 1248
Db 1259 ALDSESEKVVQOALDAAARGRTTIVIAHRLSTIQKADVIYVFDQGIKVESGTHSELVQKK 1318
QY 1249 CRYTEMCLAQSL 1260
Db 1319 CRYVELNLQSL 1330

```

## RESULT 6

US-08-612-521-2

; Sequence 2, Application US/08612521

; Patent No. 5786463

; GENERAL INFORMATION:

; APPLICANT: Peery, Robert B

; APPLICANT: Skatrud, Paul L

; APPLICANT: Thornewell, Susan J

; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF

```

; TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Division/AEH
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612.521
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamilton, Amy E
; REGISTRATION NUMBER: 33,894
; REFERENCE/DOCKET NUMBER: X-9693
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3169
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1408 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-521-2

```

Query Match 27.7%; Score 1789.5; DB 1: Length 1408;

Best Local Similarity 34.4%; Pred. No. 7.1e-178;

Matches 478; Conservative 214; Mismatches 493; Indels 203; Gaps 35;

```

QY 6 EKPRPNQDGSSESSHP-----PPKETEGRSI---SDYLIRIERYA 41
Db 92 EKPSIN---AFSKSHPIYKKSKFDFLKSRRKKKEEERKNKEKEASVLPVPSFFALFRFA 148
QY 42 DKYDWTLVNIALICATGPGCASLPLMSIIFGSGFTNKFNNY-----NSGDGSPA-----89
Db 149 APLEITANVGLVLAAGSCQPLMTLIFGRITTSFTNYAVIANQISQGLTPPTSALQ 208
QY 90 -----FKADVDFHVLWVYVLTGKFEVLTVYSTAAITISAIRTTLRVFELECTLRQEVW 144
Db 209 AAKDDLKTQSGHNALYLAIGIMFLATWLYMFWNTVGTGLNSKRIRERYLAVALROEIA 268
QY 145 HFDKQSGNGAIATXVTNNGRIQTGAELVFTVQALSMTFFSAFVVALASQWKLALITMSV 204
Db 269 YPDDLGAQEVATRIQTDCHLVQEGTSEKVALVFOYAGTFVCGFVLAIVRSPRLAGALVSI 328
QY 205 IPAIFLVTCITADAQAEARITRIYSRAAVLAERVLSSITRVHAFYAQKKWKEKYDVL 264
Db 329 LPVIMLCGGIMMTAKMYGTAAALDHAKAGSLAEVIGSIRTVQAFGKEKILGDKFADHI 388
QY 265 QQAHQEGKKKSNNGVLFSTFYFCIYAAIALAFWKGFMYQNGEADVGVKFTVVLVSVTL 324
Db 389 EQSKIVGRKSGIFEGFGLSIMFFVIYAAALAFFVGGTILVNSGO-ADSGIVINVFMSTILI 447
QY 325 AATSTMLAPSGSVVYORRIFGSELPSTIDKPTOLDPLDPSGKQPEGLGQIEIQLNLAFA 384
Db 448 GFSFMAWLAPELAANTKARGAANKLFATIDRVPAIDNSSEEGFAPDGLREISFENVKFH 507
QY 385 YPSRPSAQVLRDFNLITIPAGKTTALVAGSGSKSTMYGVLLERWYLPSPSGRILLDGLQ 444
Db 508 YPSRPSIPTLKGTFTTFFAGKTFALVAGSGSKSTVSVLSIERFYDPVSGVYVKLDGRDIRS 567
QY 445 YNVKLRSLRILVQOEPVLFRGTIFONIANGFMDQ-RDLPREKOMELVOKKACASGNDV 503
Db 568 LNLNMLRQOIGLVSOEPTLFGTTVRGVNVEHGLIGSRYNENASLEEKFEVLKKACVDANAHN 627

```

QY 504 FINEPNGYEVGERAGALSGGQRIATARSIIISDPKILLDEATSLDPAKAEKVQE 563  
Db 628 FFMKLPQGYDVMWGRGLLSGGQKQVJARAIVSDPRILLDEATSLDSEGVOD 687  
QY 564 ALNRYSKDRTLVIAHLATVKSAGNIATVISOQIVQZGTHHELI--EFGCHYAALVRAQ 621  
Db 688 ALDKASRGRTTITIAHLSTIRDAIRYVMGGEVLEQSHGNDLLANENG--PYAOLVNNQ 746  
QY 622 DL-----GADQOEHEKTLHE-----KA 639  
Db 747 KLAQAAAEALQVDDIEDPDAVFIGGSPMQEKDKQLHRAVTRGSLASIAMDDIQAKR 806  
QY 640 AREAGERPALERTHTTATSQAGLEKRVKVPVGTLYGLKILIMFYEQKNLYWCFFLLS 699  
Db 807 AEEVAGE-----DKIP---SSFGIYARLLRMNSADKFIY---ITA 840  
QY 700 TITVLICAAATFPGOALLFSRLTLVFELSG-----HAAQERADFYILMFVFWALG-----748  
Db 841 FIAAICAGWVPSLAILFGKALSDFEIQDPAELRHALSRA---LWYFITALAAAFVIF 896  
QY 749 -NLVGYFTIGTCNVISOVTHRYQAAAMFORVLDODIELLDIPQISGALTSQLSALPTQ 807  
Db 897 FOSAGFSRAGMDLNGV-----LRKKLFTATLRHDIEWFDEERNSTGAVTSNLDADQOK 949  
QY 808 LOELISANFLIYVVGORHLEQCSTTSLMMETGFGC-----VMCTSTPAF-- 853  
Db 950 VGLFGPTLGTWV-----QSCATLI-----GGCIIIGLCVGPLLALIGIACIPILVS 995  
QY 854 -GWL-----PONSRRDEARSKGLCKLCKWACKRSYRDPDRLIIDSRRPCSPVL 904  
Db 996 GYIYRLKVVVLKQEMKKLHNASAHLASAAGAVKIVASLTREK-----DVRRIYSEALK 1050  
QY 905 GHVEOGLAKIIQSFWFGFCFTHLSQSMFLAIALGFCIADVNLW---QVSTQTLNFISS 961  
Db 1051 APMKLNFRSTKS-----QCL-FAASOGIATFCIIALVFIYGA-LWIIDAKYSTASEFTVLN 1104  
QY 962 SWAFCLPQAAQYLAYSTSTFTKARSAAITLMLRTLKPTIR-ETBEENK---KKGVPVGC 1017  
Db 1105 STVFA-SIQ-AGNVFTFVPAASKANSAAISIFRSIDNEPAINAESNEGVLDHKGWVG-- 1160  
QY 1018 PVDLEDIEFRYRDRSARLVGSMTEPGQFVAYVAGSCGKSTLIALSERFYDPTSGR 1077  
Db 1161 HVRIGCVHFRPTREGVRLNLIIDVPAGTYVALVOPSGCKSTIOMLERFYDPLAGR 1220  
QY 1078 ISFAHENTAEMSPRLYRGHMLVQOEBTLYQGSVRENVTIALBAELSEELCQGRLPAR-- 1135  
Db 1221 VTLDGIDIKELNLAJSYRSQISLVSQEPTLYAGTIRFNILGANKPT-EEVTQDEIDAACK 1279  
QY 1136 -PMLWILSSLYQKALKRLAAGRMQFSGGQRIATARALIRNPKILLDEATSLDQOS 1194  
Db 1280 DANIVDTVSLPDGFTDTEVGGKSQLSGQKQRIATARALIRNPKVLLDEATSLDQOS 1339  
QY 1195 ERLVQAADDEASTRTTANVHRLSTIRNDVIVFVANGRIARTGTHAELQRLRGYEM 1254  
Db 1340 EKVQGEALDKAAKGRTTIAHRLSUSSIOHSDRIYFSEGRVADPHGTHQELLAKKGGIYEL 1399  
QY 1255 CLAQSUDQ 1262  
Db 1400 VQOMNLSR 1407

## RESULT 7

US-08-395-246C-2  
; Sequence 2, Application us/08395246C  
; Patent No. 5773214  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
; TITLE OF INVENTION: ASPERGILLUS FLAVUS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/395.246C  
; APPLICATION NUMBER: US/08/395.246C  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas G.  
; REGISTRATION NUMBER: 35784  
; REFERENCE/DOCKET NUMBER: x9683  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-2459  
; TELEFAX: 317-277-1917  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1307 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-395-246C-2

Query Match 27.6%; Score 1780; DB 1; Length 1307;

Best Local Similarity 34.3%; Pred. No. 6.1e-177;  
Matches 466; Conservative 222; Mismatches 500; Indels 172; Gaps 32;

QY 7 KPPNQDGSSE-----SHPPP---EKETEGSISYLRIFRYADKYDWTNLNVALICA 56  
Db 10 KPLKSPGTGTTTGHVSVAHEVLDRLHTPVSGIFGFIYVATRMDVAILFGSALAA 69  
QY 57 IGSGASLPMSIIFGSFTNKNFNYSNGDSPAFAKADVDHFLVWFVLYFGKFLVTVST 116  
Db 70 IAGGAALPFTVLFGRLSTFQDIATHRTVDHPEHLETKNVYFVYLYGAAEFVAILAT 129  
QY 117 AATISAITRTTLRRVLECLTQEVWHFQKSNAGATATVTTNGNRIGTGAELVFT 176  
Db 130 VGIYTDHVVQOIRVEYFQAILRONIAFFDTLGCAGEITTRITADTNLIQDISEVGLA 189  
QY 177 VOALSMFFSAFVALASQMKLALITMSVIPAFILVTGICIAIDAAQBARITRIYSRAVL 236  
Db 190 LTGLSTFVTAFTIAVTKNKKLALICSASLLALLTMTGCGSTLMLIFSKKALEYQGRASM 249  
QY 237 ABEVLSSIRTVHAFVAOKMWKDYDFLQQAHQEGKKSPNNGVLFSTEFYCYAALALA 296  
Db 250 ABEILDSIRTVAFNAQETIARKYHSLKDAEGPCKSKVIFAIWVALLCIMVNWGLG 309  
QY 297 FWKGFMVQNGEVADV--GKVTVTVSVTLAATSISMLAPSGSVVYQRRIFGSELFSID 354  
Db 310 FWMGSRFLVEG-ISNIKAGDVLITIMAILGSLYMLGNVAPNGQALSDAVAAASKLYGTID 368  
QY 355 KPTQLDPLDPSPQKPEGCLGQIEIQNLAFAPSRPSAQVLRDFNLITIPAGKTATLYGASG 414  
Db 369 QSPDLALSDQCKTLEFVRGNIVLQNRHVYPSRPEVTVAHDLSCYIPACKTATFVGP 428  
QY 415 SGKSTVWGLLERWYLPSSGRILLDGLGQVNVKWSRIELVQOEPVLPGRGTFONIAN 474  
Db 429 SGKSTIISLERFYDPVAGTILMDGHDITQTLNLRWLROOQSLVSOEPLRFTATIAENIRY 488  
QY 475 GFMEQDRLPREKOMEL--VOKACKASNGDVFINELPNGYETEVGERAGALSGGQRI 531  
Db 489 GIIGSR--PEKESTVEIRKRVAAARMANAHDFIMALPNGYDTNI--ESFSLSGGQRI 544  
QY 532 AIARSIISDPKILLDEATSLDPAKAEKVQOELNRYSKDRTLVIAHLKATVKSAGNIA 591

545 ATARAIIKDPKILLDEATSDALTSKSKLVQAALDKASKGRTTIVIAHRLSTIQKAYNII 604  
QY 592 VISOGKIVEQGHHELIIFGCHYAALVRAQDLGADEOHEKTLHEKAAREAGERPALE 651  
Db 605 VLANGOIVEQGHHELMDRGTICDMVEAHI-----KKRYSKRYSQLTNLS 655  
QY 652 RHTTAT-----SQGDL-----ERRKVPVGTGLGYS-LLKCILIM 685  
Db 656 PKHNPMTFFEDKYPGDDSDIYSLSDASDGLHTEKOR-PVSRMSLSHLMQPVKEE 714  
QY 686 FTEQKNLY-----WCLLSTIVLCAATF-PQALLPSRLTYPE-----LSGHA 731  
Db 715 AYSFWTLKFLASFNRPEWPELLGICASILAGIQPSQAVLFAKAVSTLSLPLPYPKL 774  
QY 732 QERADPYILMFVVALGMLVGYFTIGWTCNVISQVTHRYQAAAMFORVLDQDTELLDIPE 791  
Db 775 RIDANPWLMLFMIGIVSLVSYQGLFAYSSKKMYRARSQAARVILHQDISFFDQOE 834  
QY 792 QISGALTSLSALPTLOE-----LISANFL-----IYTVVGOH-----RL 827  
Db 835 NTTGALTATLSAGTKELTGISGVTGLTILVSVNLVASLGVALVIGWKLALVCISAVPAL 894  
QY 828 EOCSTLSLWMTGPGGCVWCTSTAFGWLPONSRRDEARSKLGLKCRKCRACK-RSSYR 886  
Db 895 LMGCFVRVWM-----LERFORRAKKAYQESASS-ACEAASAIR 931  
QY 887 DPDRLIFS-----RRPCSTVLGHVYEOGLAKIIQSFWRGRCFGHLSQSMFLA 936  
Db 932 TVVSLTMEALQSYOALRRQLKSDILPIVKSLL-----YASSQALPFFC 978  
QY 937 IALGFCIAVDNWL-----QVSTTOLNFTSSSWAFCLPVQAAOYLAVYSETKARSA 988  
Db 979 MALGF-----WYGSLLGHGEYSILFOYVCFSEVIF--CAQAAGVFESHAPDMGKAKIA 1030  
QY 989 ANKILWRLTKPTIRE--TENKKKG-PVGGCP--VDLEIDERYRQDSARVLRGVSMT 1043  
Db 1031 A-----REFKLFSSDMHASRSGKVPVSMRGLVEFRDVSFSPRSLEQPTLRLHNT 1084  
QY 1044 IFPGOFVAYVAGSGCKSTLIALSERFYDPTSGRISFAHENIAEMSPRLYRGHMSLVQOE 1103  
Db 1085 IKPGOFVALVAGSGCKSTTIALSERFYDPLKGGVYVDGKNIITLEMSSYRSHALISOE 1144  
QY 1104 PTLYGVSRENVTLAEAE--LSELCQGRLPARPMILWLSLYOKALKRLAAQRCMQFS 1161  
Db 1145 PTLFOGTIRENILLGNTPHVTDDFLVKACKDANIYDFILS--LPOGFNTIVNGKGMLS 1202  
QY 1162 GGQORAIARALINPKILLDEATSDALTSQSERLVQAALDEASTSRTTIAVAHRLSTI 1221  
Db 1203 GGQORAIARALINPKILLDEATSDALTSQSEKVVQAALDAAARGRTTIAVAHRLSTI 1262  
QY 1222 RNVDFVIFVANGRTAETGTHAELQRLRGRYEMCLAQSLD 1261  
Db 1263 QRADLIYLDQGEVVEVSETHRELLKKGRYVELVHLQNP 1302

RESULT 8  
5206352-4  
; Patent No. 5206352  
; Applicant: Roninson, Igor B.; Pastan Ira H.; Gottesman,  
; Michael M.  
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA  
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
; NUMBER OF SEQUENCES: 4  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/622,836  
; FILING DATE: 24-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 892,575  
; FILING DATE: 01-AUG-1986  
; APPLICATION NUMBER: 845,610  
; FILING DATE: 28-MAR-1986  
; SEQ ID NO: 4:  
; LENGTH: 1280

5206352-4

Query Match 25.4%; Score 1640.5; DB 6; Length 1280;  
Best Local Similarity 32.3%; Pred. No. 2.7e-162;  
Matches 427; Conservative 243; Mismatches 491; Indels 163; Gaps 26;

QY 24 EKTEGSDYLRIFRYADKYDMLTNVIALICAIAGSGASLPLMSIIFGFTNKNFN----- 79  
Db 27 KKEKKPTVSF-SMFRYSNKLKLYMVGTLAAIIGHAGCLPLMLVFGEMTDIFANAGNL 85  
QY 80 -----YNSGDSPPAF-----KAUDHVLVFWLFTGKFLVTVVSTAAITISAIRTR 128  
Db 86 EDMSNTNRSNDYDGFPMNLEEDMTRYAYISGIGAGVLVAAVQVSWCLAGAGROIH 145  
QY 129 TLRVFLLECTLROEVWHFDKQSGAIGATVTTNGNRITGTGAELKLVFTVQALSFFSFAV 188  
Db 146 KIRQPFHAIMROEIGWFDVHVGELNTRLTDDVSKINQVIGDKIGMFFQSNATFFTGFI 205  
QY 189 VALASOWKLALITMSVIPALFLVTGICIAIDAAQEARITRYSRAVLAEEVLSIRTVH 248  
Db 206 VGTGRGKLTVLILAIISPVGLSAAVWAKILSFTDKELLAYAKAGAVEEVLAAIRTVI 265  
QY 249 AFTAQKKWKEKYDYFLOQAHOQKXKSPNGVLFTSTFYCYAATAAFWKGPRMYQNGE 308  
Db 266 AFGQKKELERYKNLEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLVSGE 325  
QY 309 VADVGVFTVVLSTLAATSIISMLAPSGSVYQRRIFGSELSIIDKPTQLDPLDPSGKQ 368  
Db 326 YS-IGOVLTVPFVSLIGAFSGVQASPSIEAFANARGAAVEIFKIIDNKPISDSYKSGHK 384  
QY 369 PEGCLGOIEIQNLAFAYPRPSAOVLDRDNLIPACKTALVAGSGCKSTMVGLLEHWY 428  
Db 385 PDNIKMLERNVHFSPSRKEVKILGLNLKVGSGQTVLGVNSGCCSKSTVOLMORLY 444  
QY 429 LPSSGRILLDGLBGQYNYVKMLSRILVQOEPVLFRGTIFONTIANGPMDQORDLPREK- 487  
Db 445 DPTGMSVSDGDIRINVRFLREIGVVSQEPVLFAFYATIAENIRYG-----RENV 495  
QY 488 QMELVOKACKASGDVFINELPNGVETEVERAGALSGQGRQRIARSIISDPKILLD 547  
Db 496 TMDEIEKAVKEANAYDFINKLPHKFDTLVGERGAQLSGGOKORIAARALVRNPKILLD 555  
QY 548 EATSALDPRAEKVVOBALNRVSKDRDTLLIAHKLATVKSAGNIAVISOCKIVEGTHHEL 607  
Db 556 EATSALDTESEAVVQVALDKARRGRTTIVIAHRLSTVRNADVAGFDGDIVERKGNDEL 615  
QY 608 IFPGCHYAALVRAQDLGADEOQOHEKTLHEKAAREAGERPALERT----- 653  
Db 616 MKEKGYFLKLVMTQTAGNEVELEN-----AADESKSEIDALEMSSNDRSSSLIRKST 668  
QY 654 -HTTATSOAGD-----LEKRVKVPVGTGLGYSLLKCLILIMFYEQKNLYM-CFLLSITIV 703  
Db 669 RRSVRGSAQDRKLSLTKALDESIPVSF--WRINKNL-----TENYFVVGVFCA 718  
QY 704 LICAATFPQALLPSRLTYPEL--SGHAAQERADFYILMFVVALGNLVGYFTIGWTCN 761  
Db 719 IINGGLOPAIIFIISKIIGVTRIDDPETKRONSLFLLFLAIGIISFTIFFGQGTFF 776  
QY 762 VISQVTHRYQAAMFORVLDQDTELLDPEQISGALTSLSALPTLOQELISANFLIYIV 821  
Db 779 KAGEILTCLRKYVWFRSMLRQDVSWFDOPKNTGALTALRDLANDAAQVKAIGSLAV--- 835  
QY 822 VGOHRLQECSTLSLWMTGPGGCVWCTSTPAFGW-----LPO 858  
Db 836 -----IFONTIANLGTGIII-----SFYVQWQLTLLLAIVPITAIAGVVEKMLSG 881  
QY 859 NSSRDEARSKLGLKCRKCRKCRKCRKCRKCRKCRKCRKCRKCRKCRKCRKCRKCRK 916  
Db 882 QALKDKKELEGACKIATEA-----IENFRTVVSLSLTOEOKFEHMYAQSLQVP 927  
QY 917 -----SFWFGRCFCFHLSQSMFLAIA-----LGFCITAVDNWLVQVSTTQLNFISSSWAF 965



Db 983 PGAMAVQVSSFPADYAKAKISAAHIIIMIENTPLIDSYSTE-----GLMPTNTEGN 1034  
Qy 1023 ----DIEFRYQRDSARVLGVSMTEPQGFVAYYGASGCGKSTLIALSERFYDPTSGRI 1078  
Db 1035 VTFGEVVFNTPTRPDIPVLQGLSLVKKGTTLALVSGSGCKSTVVQLLERYDPLAGKV 1094  
Qy 1079 SFAHENAEMSPRLYRGHMSLVQOEPYLVGSRVENVTLALEAEI--SEELCOGRLPARP 1136  
Db 1095 LLDCKEIKLVNQLWRAHLGIYVSOEPLDFDCSIAENIAYGNSRVVSOEIVRAAKANI 1154  
Qy 1137 MLWTLSSLYQKALKRAAQEGWFGSGGQQRQRTAIALIRPKLILLDEATSDALDTOSER 1196  
Db 1155 HAFTESLPNKYSTK--VGDKGTQLSGGQKQRTAIALIRPKLILLDEATSDALDTSEK 1212  
Qy 1197 LVQAALDEASTRTTIAVAHRLSTIRNDVIEVFANGRIAEITGHAEIOLRLGRYYEMCL 1256  
Db 1213 VYQALDKAREGTCIVIAHRLSTIQNDLIIVFONGRVEHGTQQLAQKGIYFSMWS 1272  
Qy 1257 AQS 1259  
Db 1273 VQA 1275

## RESULT 10

US-08-784-649A-2  
; Sequence 2, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; APPLICANT: Chen, Gang  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: Reg. No. 5830697 36,677  
; REFERENCE/DOCKET NUMBER: 06037/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5070  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1279 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-784-649A-2

Query Match 25.3%; Score 1631; DB 2; Length 1279;  
Best Local Similarity 32.3%; Pred. No. 2.7e-161;  
Matches 427; Conservative 243; Mismatches 490; Indels 164; Gaps 27;  
Qy 24 EKEGSGISDYLRIFRYDKYDWTNLVIALICALIGSGASLPLMSIIFGSPFNKFN-- 79  
Db 27 KKEKKFTVSVF-SMFRYSNWLKLYVWVYVTLAAIHGAGLPLMLVFGSEMTDIFANAGNL 85

Qy 80 -----YNSGDGSPPEAF-----KADVDHVFYLMFVYLFIGKFLVTVYSTAAITISAIRTR 128  
Db 86 EDLMSNTNRSIDINDTGFEMNLEEDTRVAYYSGIGAGLVAAVYQVSEMFCAAGROIH 145  
Qy 129 TLRVFELECTLRQVWHFDKQSNCAJATVTTNGNRIOGTIAEKLVFTQVALSMFEFAFV 188  
Db 146 KIRQGFHAIHROEIGHFDFVHDVGNLTRLTDDVSKINEVIGDKICMFPQSWATFTGFI 205  
Qy 189 VALASQWKLALITMSVTPAIFLVGTICIAIDAAQEARITRIYSRAVLAEEVSSIRVH 248  
Db 206 VGFTGKRLTLVILATSPVLGSLAAVWAKTLSSFTDKELLAYAKAGAAVEVLAIRTVI 265  
Qy 249 AFYAQKMKVYDVFLQAHQEGKSPNGVLFSTFEYECIYAAALAFWAKGFRMYONGE 308  
Db 266 AFGQKKELEYRKNLPEAKRIGIKKAITANISIGAAFLIYASVALAFWYGTTLVLSGE 325  
Qy 309 VADYGVFTVVLVSLAATSISMLAPSGSVYVORRIFGSELSIITDKPTQLDPLDPGSKQ 368  
Db 326 YS-IGQVLT-VFSLIGAFSVGOASPSLEAFANAKAAVEIKFIIDNKPSIDSYKSGHK 383  
Qy 369 PEGCLGOIEIONLAFAYPSRPSAQLVDFNLIPACKTALVAGSGSGKSTWGLLERY 428  
Db 384 PDNTKGNLEFRNVHFSYPSRKEVKILKGLNLKAVQSGQTVLVAGSGCGKSTVQLMORLY 443  
Qy 429 LPSGRILLDGLGQYVNVKWLRSIRLVQOEPLVFRGTIFQNIANGFMQORDLPREK- 487  
Db 444 DPTGMYSDVGDQDITINVFLREIIGVVSQEPVLPFATTIAENIRYG-----RENV 494  
Qy 488 OMELVQKACKASNGDVFINELPNGYETEGERAGALSGGQRIARISITSDPKILLD 547  
Db 495 TMDEIKAKVKNAYDFIMLPHKFDTLVGERCAQLSGGQKORAIARALVRNPKILLD 554  
Qy 548 EATSALDPKAEKVQOALNRVSKDRTILVIAHKLATVKSAGNIAVISOCKIVEQGTTHLE 607  
Db 555 EATSALDTESEAVVQALDKARKGRTIIVIAHRLSTVRNADVIAGFDGIVIVEGNHDEL 614  
Qy 608 IEFQCHYAALVRAQDLGADQEQHEKTLHEKAAREAAGERPALERT----- 653  
Db 615 MKERGIYFKLVMTQTAGNEVELEN-----ANDESKSEIDALENSNDRSSIRKRST 667  
Qy 654 -HTTATSOAGD-----LEKRPVWPVGLYSLKLCILIMEYOKLYW-CFLLSTITV 703  
Db 668 RRSVRGSOAQDKLSTKEALDESIPVSP--WRINKNL-----TEMPYFVGVPCA 717  
Qy 704 LICAAETPGQALLPSRLTYVEL--SGHAAQERADFYILMEFVVALGNLVGYFTIGTCN 761  
Db 718 IINGCLOPAFIIFSKIIGVFTRIDDPETKRONSLFLLFLALGIIISFTIFFLOGFTFG 777  
Qy 762 VISQVTHRYQAAMFORVILDDIPEIQSGALTQSLSALPTQLOELISANFLIYIV 821  
Db 778 KAGEILTKRLRYMVFRLSRQDVSFWDDPKNTTGALTTLRLANDAAQVKAIGSLAV--- 834  
Qy 822 VGOHRLBQCSTTSLMMETPGGCVNCTSTPAFCW-----LPQ 858  
Db 835 -----ITQNIANLGTGIII-----SFIYQWLTLLLLAIVPIAIAAGVWENKMLSG 880  
Qy 859 NSSRDEARSRLKLCRCWKACKRSSYRDPDLIFDSRRPCSPVTLGHVGOGLAKITQ-- 916  
Db 881 QALKDKKELEGACKIATEA-----IENFTVVSILTQEQKFEHMYAQSLQVP 926  
Qy 917 -----SFWFGRCFGFHLQSMEFLAIA-----LOFCTAVDNWLOVSTQNFNLISSSWAF 965  
Db 927 YRNSLRKAHIFGTTF--TOAMMYFSYAGCFRFGAYLVAHKLMSFEDVLVFS-----AV 980  
Qy 966 CLPVOAAQYLAYSTFTTKARSAANYILMLRLTKPTI--RETEENKKKPGVGGCPVDLE- 1022  
Db 981 VFGANAVQGVSSFPADYAKAKISAAHIIIMIENTPLIDSYSTE-----GLMPTNTEGN 1034  
Qy 1023 ----DIEFRYQRDSARVLGVSMTEPQGFVAYYGASGCGKSTLIALSERFYDPTSGRI 1078  
Db 1035 VTFGEVVFNTPTRPDIPVLQGLSLVKKGTTLALVSGSGCKSTVVQLLERYDPLAGKV 1094









CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,537

FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas G.  
REGISTRATION NUMBER: 35 784

REGISTRATION NUMBER: 35,764  
REFERENCE/DOCKET NUMBER: X9212  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-2459  
TELEFAX: 317-276-1917

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1202

LENGTH: 1302 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
1-232-537-2

Query Match 20.3%; Score 1308.5; DB 1; Length 1302;  
Best Local Similarity 29.3%; Pred. NO. 2.3e-127;  
Matches 379; Conservative 245; Mismatches 549; Indels 119; Gaps 33;

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 791 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-394-880B-2

Query Match      10.0% Score 647.5; DB 1; Length 791;
Best Local Similarity 31.5% Pred. No. 2.8e-58;
Matches 188; Conservative 107; Mismatches 231; Indels 71; Gaps 17;

QY 46 WTLNVIA-----LCAIGSGASLPLMSIIF--GSFTNKFNNYNSGDGSPFAFKA 92
DB 168 WRLLIARPEAKKIALAFLFLVSSGIT--MSIPFSIGKIMPTSTKATT-EGGNELFGL 223
QY 93 DVDHEVLMFVYLFIGKF--VLT-----YVSTAAITISAIRTTTLRRRVFLECTLRQEVW 144
DB 224 SLP-----MFYGALAGILTLGAAANYGRIIILRIVGERIVARLRSKLFRTFVQDAE 275
QY 145 HDKOSNGAIATXVTN---GNRIQTGIAEKLVFVQALSMFFSAFVVALASOWKLALI 200
DB 276 FDANRVGDLSRLSSDTIIVGKSITQNLSDGLRAAVSGAAGFGLMAYVSL----KLSSI 331
QY 201 TMSVIPALFLVTGICIAIDAAQEARITRIYSR-----AAYLABEVLSSIRTVHAFYA 252
DB 332 LALLLP-----IGLAFYGYGRAIRNLSROIQRNLGTLTKIAERLGNVKTSSQSFAG 383
QY 253 QKKMVEKYDVLQQAQHEGKKKSPNGVLFSTEFECIYAAIALAFWKGFMYONGEVADY 312
DB 384 EVLEVRYNQVKRIFELGKRRESLISATFFSSTGFAGNMTILALLYGGWVQSGAIT-I 442
QY 313 GKVFTVLSVTLAATSIAPSGSVVYQRRIFGSELFSDIKPTQLDPLDPGKQPEGC 372
DB 443 GELTSFLMYTAYAGSMFGLSSFYSELMKGVGAASRLFELQDRQPTISP--TKGEKVASA 500
QY 373 LGOIETIONLAFYSPRPSAOVLDFNLTPAGKTTALVAGSGGSKSTMVGLLERWYLPSS 432
DB 501 RGFIRFENVTFSPTRPAVPFIRDLNFEIPQGTNVAIVGSPGGGKSTIASILLRFYSPT 560
QY 433 GRILLDGLGQYNNVKYLRSLRVLVQEPVLFRCGTFPONTANGFMDEQRDLPREKOMELY 492
DB 561 GRVLGGKDITHMNAKSLRRKIGIVSQEPVLFSGTIAENTAYG-----KQAKRSEIV 613
QY 493 QKACKASNGDVFINELPNGYETEVGERAGALSGGQRIARSIISDPKILLIDDEATSA 552
DB 614 AAARKAN--CQISDFPDGLDTQVGRGAQLSGGQKQRIARALIKDPDILIDDEATSA 671
QY 553 LDPKAEKVVOEALNRVSK-DRTTLVIAHKLATYKSNAGNIAVIS-OGKIVEQGTTHHEL 607
DB 672 LDAESETLVNSALTALLRGNNTTISIAHRIJSTIKRSDTIIVLPGDGRVAFQGSYEEL 728
```

Search completed: April 27, 2002, 07:47:09  
Job time: 6726 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2002, 07:45:35 ; Search time 104.96 seconds  
(without alignments)  
1336.568 Million cell updates/sec

Title: US-09-882-694A-11

Perfect score: 6455

Sequence: 1 MADESEKPPNQDSESSSH.....LQRLRGYYEMCLAQSLDQA 1263

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6453	100.0	1263	22	AA1980
2	2392	37.1	1248	22	AA1981
3	2270	35.2	1308	20	AA1982
4	1956	30.3	1349	20	AA1983
5	1940	30.1	1334	20	AA1984
6	1934.5	30.0	1333	22	AA1985
7	1789.5	27.7	1408	17	AA1986
8	1780	27.6	1307	17	AA1987
9	1656	25.7	1283	22	AA1988
10	1654.5	25.6	1280	22	AA1989
11	1649.5	25.6	1280	18	AA1990

12	1649.5	25.6	1280	21	AA1991
13	1649.5	25.6	1280	22	AA1992
14	1647.5	25.5	1280	11	AA1993
15	1647.5	25.5	1280	22	AA1994
16	1643.5	25.5	1280	22	AA1995
17	1643.5	25.5	1280	22	AA1996
18	1641.5	25.4	1272	21	AA1997
19	1641.5	25.4	1281	22	AA1998
20	1640.5	25.4	1281	22	AA1999
21	1640.5	25.4	1272	21	AA2000
22	1640.5	25.4	1280	8	AA2001
23	1640.5	25.4	1280	21	AA2002
24	1640.5	25.4	1280	22	AA2003
25	1640.5	25.4	1280	22	AA2004
26	1640.5	25.4	1281	22	AA2005
27	1640	25.4	1279	22	AA2006
28	1638.5	25.4	1279	22	AA2007
29	1636.5	25.4	1280	14	AA2008
30	1636.5	25.4	1281	15	AA2009
31	1635.5	25.3	1280	22	AA2010
32	1631	25.3	1279	19	AA2011
33	1630.5	25.3	1276	14	AA2012
34	1630.5	25.3	1276	21	AA2013
35	1629.5	25.2	1276	21	AA2014
36	1629.5	25.2	1280	19	AA2015
37	1623.5	25.2	1280	19	AA2016
38	1623.5	25.2	1280	19	AA2017
39	1617.5	25.1	1280	19	AA2018
40	1601	24.8	1275	21	AA2019
41	1601	24.8	1275	22	AA2020
42	1589	24.6	1261	22	AA2021
43	1573	24.4	1313	22	AA2022
44	1507	23.3	1275	20	AA2023
45	1495	23.2	1286	21	AA2024

#### ALIGNMENTS

RESULT 1

AA1980: AAY72637 standard; Protein: 1263 AA.

AC AAY72637;

XX 02-MAY-2001 (first entry)

DT Exophiala spinifera p-glycoprotein, a fumonisin degradative enzyme.

XX Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;

XX aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;

XX detoxification; mycotoxin; animal feed; human feed; sludge;

XX transgenic plant; transgenic animal; microbial spray.

XX Exophiala spinifera.

XX Key Location/Qualifiers

XX Misc-difference 157

XX /label= Unknown

XX /note= "Encoded by CAR"

XX WO200105980-A1.

XX 25-JAN-2001.

XX 14-JUL-1999; 99WO-US15824.

XX 14-JUL-1999; 99WO-US15824.

XX (PION-) PIONEER HI-BRED INT INC.

XX (CURA-) CURAGEN CORP.

XX Duvick JP, Maddox J, Gilliam J, Folkerts O, Crasta OR;

XX WPI; 2001-147345/15.  
DR N-PSDB; AAD02694, AAD02695.  
XX Novel polynucleotides encoding Exophiala degradative or transport  
PT enzyme which is useful for detoxifying fumonisin or structurally  
PT related mycotoxin during processing of grain for human or animal food  
XX consumption -  
PS Claim 1c: Page 77-82; 90pp; English.  
XX The patent discloses novel polynucleotides encoding Exophiala spinifera  
CC fumonisin degradative or transport enzymes such as flavin monooxygenase,  
CC aldehyde dehydrogenase, permease and p-glycoprotein. A fumonisin esterase  
CC enzyme with at least one fumonisin degradative enzyme is useful for  
CC detoxifying fumonisin or a structurally related mycotoxin present in  
CC harvested grain, on application to a plant or to harvested grain  
CC during processing, or to processed grain that is to be used as animal  
CC or human feed, or as a silage. Nucleotide fragments of the present  
CC invention are useful as probes and primers. They can be introduced  
CC into microorganisms that multiply on plants to deliver enzymes to  
CC potential target crops. The genes encoding the degrading enzymes are  
CC introduced via a vector into a microbial host and the transformed host  
CC is supplied to the environment, plants or animals for reducing the  
CC pathogenicity of a fungus producing fumonisin. The genes of the  
CC invention are fermented in a bacterial host and the resulting bacteria  
CC is processed and used as a microbial spray. The nucleotide sequences  
CC can be used alone or in combination to engineer microbes or other  
CC organisms to metabolise fumonisin and resist its toxic effects.  
CC The present protein sequence is p-glycoprotein, a fumonisin  
CC degradative enzyme from Exophiala spinifera.  
XX Sequence 1263 AA;  
SQ

Query Match 100.0%; Score 6453; DB 22; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADESEKPRPNQDSSESSHPPEKETEGSISDYLRFRVADKYDWTNLNVIALICAISG 60  
DB 1 madesekprpnqdsessshpppeketegsisdyrlfrvadydwtlnvialicaisg 60  
QY 61 ASPLMSIIFGSFNMKNVNSGDSPEAFKADVDHFLVFLVFCFKVLVYVSTAAT 120  
DB 61 asplmsiifgsfntknfnynsgdspeafkadvdhfvlvflvfcfkvlvystaat 120  
QY 121 ISAITTTTLRRVLECLTRQEVWHFDKQNGAINTXVTNGNRITGTGAELKLVFTVQAL 180  
DB 121 isaitttrtlrrvlecltrqevwhfdkqngaintxvtngnrigtglaelklvftvqal 180  
QY 181 SMFSAFYVALASQWKLALITMSVIPALFLVTGICIAIDAAQEARITRIYSRAAVLAREV 240  
DB 181 smfafaavlasqwkalltmsvipalfvltgicidaaqaearitriysraavlaev 240  
QY 241 LSSIRTVHAFYAKMKVEYDVFLOQAHOEGKKSNNGVLESTVEYFCIYAALAFWKG 300  
DB 241 lssirtvhafyakmvekydvflqqahogekksnngvlesteveyfciaalafwkg 300  
QY 301 FRMYONGEVADGVFTVVLVSVTLAFTSISMLAPSGSVYQRRIFGSELSIIDKPTQID 360  
DB 301 frmyongevadgvftvvlsvtlaftsismlapsgsvvyqrrifgselsiidkptqid 360  
QY 361 PLDPSGKQEGCLQIETQNLAFAYPSPSAQVLRPNLTIPAGTKTTLVAGSGSGKSTM 420  
DB 361 pldspgkqegclqieqlnafaypspsaqvldrpnltipagtkttlvagsgsgkst 420  
QY 421 VCLLERWLPSSGRILLDGLGQYVYKWLRSRIRLVQEPVLFRTQNTANGFMEQ 480  
DB 421 vcllerwlpssgrilldglgqyvvykwlrsrirlvqepvlftrtqntangfmeq 480  
QY 481 RDLPREKOMELVQKACKASNGDVFINELPNGETEYGERAGALSGGQRQRIARSIIISD 540  
DB 481 rdlprekomelvqkackasngdvfinelpngeteeygeragalsggqrqriarsiiisd 540

DB 481 rdlprekqmelvqkackasngdvfinelpngyetevegeragalsggqrqriarsiiisd 540  
QY 541 PKTILLDEATSAIDPRAEKVQDEALNRVSKDRTTLVIAHKLATVKSAGNIAVISOGKIVE 600  
DB 541 pkillldeatsaidpraekvqdealnrvskdrttlviahklatvksagniavisogkive 600  
QY 601 QGTHHELIEFGCIYAAALVRAQDLGADQDEQHEKTLHEKAAREAAERPALERTHTTATSQ 660  
DB 601 qgthheliefgciyaaalvraqdlgadegdehaktlhekareaaerpalerthttatsq 660  
QY 661 AGDLERKRVPGVGLYSLLKCILIMFYEQNLVWCFLLSTITVLITCAATFPQALLFSRL 720  
DB 661 agdlekrkvpvglysllyllkclilimfyeknlwcfllstltvlltcaatfpgallfarl 720  
QY 721 LTVFELISCHAAQADRFYILMFVVALGNLVGYFTIGWTCNVISQVTVRVRQAAMFQRLV 780  
DB 721 ltvfelischaqaadrfyilmffvvalgnlvgyftigwtcnvvisqvvtvrrvraqamfgrvl 780  
QY 781 DDITELLDIPEQISSGALTSQSALPTQLOELISANFLIYVVGQHRLEQCSSTTSLMETG 840  
DB 781 ddiellldipeqissgaltsgsalptqlqelisanfliyvvgqhrlegcsttslmetg 840  
QY 841 PGCYWCCTSTPATGWLPPQNSSRDEARSKLGLCRKCKACKRSSYRDPDRIFDSRRPCS 900  
DB 841 pgcywctstpatgwlppqnssrdearsrklglckrckackrssyrdpdrilfdrpccs 900  
QY 901 PTVLGHVEQGLAKITIOSFWFGKCFPHLSQSWEFIAIALGFCIAVDNWLQVSTTQLNPI 960  
DB 901 ptvlghveqglakitiosfwfgkcfphlsqswefialalgfciavdnwlqvsttqlnpi 960  
QY 961 SSWAFCLPVOAAQYLAYSTSTTKARSAANYILWLRTLKPTIRETEENKKGPGVGGCPVD 1020  
DB 961 sswafclpvqaaaylayststkarsaanyilwlrtlkptireteenkkgpgvggcpvd 1020  
QY 1021 LEDIERVRRORSARVLRGVSMITIEPGOPVAVVGASGCKSTLIALSERFVDPVTSGRISF 1080  
DB 1021 lediefrrvrrorsarvlgvsmitiepgopvavvgasgckstlialserfvdvptsgrisf 1080  
QY 1081 AHENIEMSPRLYRGHMSLVQOEPTLYGQSVRENTVLALEAFSELSCQGRIPARPMIWI 1140  
DB 1081 aheniemspriyrghmslvqqeptlyqgsvrenvtlaaleafselscqgriparpmiwi 1140  
QY 1141 LSSLYQALKRLAAQRMQFSGGQRQRIARALIRNPKLILLDEATSAIDQTSERLVQA 1200  
DB 1141 lsalyqalkrlaaqrmqfsggqrqriaralirnpkllldeatsaidqtserlvqa 1200  
QY 1201 ALDEASTSRITTAHARLSTIRNVDFIVFPFANGRIAEFTGTHAEQRLRGRIYEMCLAQSL 1260  
DB 1201 aldeastsrlltavahrllstirnvdfivfpfangriaetgthaeqlrgriyemclaqsl 1260  
QY 1261 DQA 1263  
DB 1261 dqa 1263  
RESULT 2  
AAB62488  
ID AAB62488 standard: Protein; 1248 AA.  
XX  
AC AAB62488;  
XX AC  
DT 09-JUL-2001 (first entry)  
XX  
DE P. chrysogenum ABC transporter aal sequence.  
XX  
KW Beta-lactam; micro-organism; ABC transporter; ATP-binding cassette;  
KW adenosine triphosphate-binding cassette; antibiotic; penicillin;  
XX cephalosporin; aal.  
OS Penicillium chrysogenum.  
XX  
PN W0200132904-A2.  
XX



PD 10-MAY-2001.  
 PF 03-NOV-2000; 2000WO-EPI1489.  
 XX 03-NOV-1999; 99EP-0203684.  
 PR 03-NOV-1999; 99EP-0203685.  
 PR 03-NOV-1999; 99EP-0203687.  
 PR 03-NOV-1999; 99EP-0203688.  
 PR 03-NOV-1999; 99EP-0203689.  
 PR 03-NOV-1999; 99EP-0203690.  
 PR 03-NOV-1999; 99EP-0203691.  
 PR 03-NOV-1999; 99EP-0203692.  
 PR 03-NOV-1999; 99EP-0203693.  
 PR 03-NOV-1999; 99EP-0203694.  
 XX (STAM ) DSM NV.  
 XX Van Den Berg MA, Bovenberg RAL, Driessen AJM, Konings WN;  
 PI Schuurs TA, Nieboer M, Westerlaken I;  
 XX N-PSDB; AAF83380, AAF83381.  
 DR WPI; 2001-291055/30.  
 DR Enhancing secretion of beta-lactam compounds from a micro-organism by  
 PT enhancing adenosine triphosphate-binding cassette transporter activity,  
 PT useful for producing e.g. penicillin and cephalosporins -  
 XX  
 XX Claim 8; Page 56-58; 116pp; English.  
 XX The invention relates to a method for enhancing the secretion of beta-  
 CC lactam compounds from a micro-organism that comprises enhancing ABC  
 CC (adenosine triphosphate (ATP)-binding cassette) transporter activity of  
 CC the micro-organism. The method is used for enhancing the production and  
 CC secretions of beta-lactam antibiotics like penicillin and cephalosporins.  
 CC Manipulation of the ABC-transport protein system provides a means for  
 CC enhancing beta-lactam secretion. The present sequence represents the  
 CC P. chrysogenum ABC transporter aal protein sequence.  
 XX  
 XX Sequence 1248 AA;

Query Watch 37.1%; Score 2392; DB 22; Length 1248;  
 Best Local Similarity 41.3%; Pred. No. 1.7e-214;  
 Matches 551; Conservative 205; Mismatches 417; Indels 160; Gaps 25;

Qy 1 MADESEKPRNQDSSESSH-----PPPEKETEGESISDYIRIFRYADKYDWTLNVI 51  
 Db 1 mekesvkysdgggaeeqtnafngylvappplaglikanirk--riifqygsakvymqgi 58  
 Qy 52 ALICATGSGASLPLMSIFGSEFTNKNFNYSNGDGPFAFKADVDHFLVFLYFIFGRFVL 111  
 Db 59 atacalgsgvmaInivfgfittvdytsgridparfrgeagrsllyffvlglerfll 118  
 Qy 112 TVVSTAATTAITRTTLRVECTLRQEVVHFDKOSNGAIAATXVTNNGNRTQICIAE 171  
 Db 119 tyyystlftiaayritnirlylkagltqelaiffdsagsgsiavqatsngrliqsgise 178  
 Qy 172 KLVTFTQALSMFFSFAVVALASQWKALITMSVIPAFILVVICICIAIDAAQEARIRIYS 231  
 Db 179 klgvlvqqlstfisarvlarfvhkwltlccclapamllalgissaleasletkvkva 238  
 Qy 232 RAAVLAEEVLSIRTHAFYAQKMKVYDFVLQQAHOEGKKSPNNGVLFSTEYFCIYA 291  
 Db 239 qagafaeslssstrvqafgrkqlldfrdrlqasrnlgnkksplfgclfatayslifa 298  
 Qy 292 ATALAFWKGFRMONGVADVGVKVTVLVSVTLAATSISMLAPSGSVVYQRRIFSGELFS 351  
 Db 299 gfglcfwgikxmvstgevaqsgdiftvmsvvaatsltsilpyvidfvraasaaelfq 358  
 Qy 352 IIDKPTOLDPDSKQPEGLQIEIQNLAFAYPSRQAQVLRDFNLTIPAGTTPALVG 411  
 Db 359 lmdrtsrldpdesgerpdcvtvdydfggsifayrprpdkvindhfsihupagkttalvg 418

Qy 412 ASGSKSTMVGLLEWYLPSSGRITLLDGLGELGOYNVKKWLSRIRLVQOEPVLPFGTIFQN 471  
 Db 419 asgskstivglrierwydplsgtikidggtietfnikwlrqgvrlvgqepvlfagvaen 478  
 Qy 472 IANGFME--ORDLPREKQNELVOKACKASNGDVFINELPNGETEVETEYEGAGALSQGRQ 529  
 Db 479 ianglvgtptwenesaaekar-igdaakiafahdfitelpngydvtvigerggllsgqkq 537  
 Qy 530 RIAIARSIIIDPKILLDEATSDALPKAEKVQOELNRYSKDRTTTLVIAHLKLVKSGN 589  
 Db 538 rvaiaarsivsqprililideatsalidphaeevvgralnnvskgrttititiahkirtidahn 597  
 Qy 590 IAVISOGKIVGEGTHHELIEFGCHYAALVRAODL-----GADE-QOEHEKTLH 636  
 Db 598 lvmengrileqgthksllidngayarlvaqdisvaacnqpedvsnssdetkdehvtlt 657  
 Qy 637 EKAAREAGERPALERTHTTATSQAGDLERKRVVGTGLGYSLKLCILIMFYEQKNLYWCF 696  
 Db 658 geltrcstmrtsalek-----rlsrdddnwk-----rlg--llhtwrlvskstpelntwf 706  
 Qy 697 LLSTITVLIATATPGQALLFSRLTLVFLSGHAAQERADFYILMFFVVALGNLVGYFTI 756  
 Db 707 aalliscilageasfpgailmsrfievftqgaamrtkgnffalmflvlgatvvyfvv 766  
 Qy 757 GWTCNVISQVTHRYQAAAMFORVLDODIELDIPQISGALTSQLSALPTOLOELISANF 816  
 Db 767 gwtcnvvvqumhkyrkqvndmmrqdlqffdraenttgalt-----pt----- 811  
 Qy 817 LYIVVGHURLBOCSTTSLMMETGGCVWCTSTPAFGW-----LPONSREAR 866  
 Db 812 -----vg-----viscvial-----ayawklglvllaglpmlisgyar 847  
 Qy 867 SRKLGKLRK-----CWACKRSS-----YRDPRLIFDSRRPC 899  
 Db 848 lrmegamdhkiskmcgsasiasaevaitrtvsalaetsvlerytckeldranatstkp 907  
 Qy 900 SPTVLGHVBOGLAKIIQSFWRGCFPHLSQSMELAIAGFCIAVDNMLQV-----ST 953  
 Db 908 lllml-----pfafqtveysflalgf-----wygcrivsgdv 941  
 Qy 954 TOLNITSSWAFCLPVQAAQYLAYSTFTKARSAANYILMLTLKPTIREENKKGP 1013  
 Db 942 smvnffislvfvggqatillfgsskmhskp-----vw--eakptiretdeardigp 994  
 Qy 1014 VGGCPVDLEDIEFRYRDRDSARVLGVSMTEPGQVAVTVGASGCGKSTLIALSERYPDP 1073  
 Db 995 gdlkslhleslqfyspmrpharvlrgidihlekeqivafvgasgcgkatmslierfydp 1054  
 Qy 1074 TSGRISFAHENTAEPSRLYRGHMSLVQOEBPTLYOGSVRENVTLAEALSELSEELCOGRLP 1133  
 Db 1055 vaghikidaltidslnpwlryggvalvqeplypgtirenvamgtptddgstvpsdie 1114  
 Qy 1134 A---RPMWLSSLYOKALKRLAAQRMQFSGGQRIARALRNPKLLILDDEATSA 1190  
 Db 1115 aacraanawefisslpdglmlcngangtqlsggqrgriarallnrprllildeatsal 1174  
 Qy 1191 DQSERLVOALDEAST--SRTTAVAHRLSTIRNVDFVFANGRIATGTTHALQELR 1248  
 Db 1175 dtqserlvqdaineaahydriftiavahrlistirhadmicvfdggriaesghehel-lmk 1233  
 Qy 1249 GRYV-EMCIAQSL 1260  
 Db 1234 grylpkmeaql 1246  
 RESULT 3  
 ID RAY21815  
 XX RAY21815 standard; Protein; 1308 AA.  
 AC RAY21815;  
 XX  
 DT 10-SEP-1999. (first entry)  
 XX







XX 10-MAY-2001.  
 XX PD 03-NOV-2000; 2000MO-EP11489.  
 XX PF 03-NOV-1999; 99EP-0203684.  
 XX PR 03-NOV-1999; 99EP-0203685.  
 XX PR 03-NOV-1999; 99EP-0203687.  
 XX PR 03-NOV-1999; 99EP-0203688.  
 XX PR 03-NOV-1999; 99EP-0203689.  
 XX PR 03-NOV-1999; 99EP-0203690.  
 XX PR 03-NOV-1999; 99EP-0203691.  
 XX PR 03-NOV-1999; 99EP-0203692.  
 XX PR 03-NOV-1999; 99EP-0203693.  
 XX PR 03-NOV-1999; 99EP-0203694.  
 XX PA (STAM ) DSM NV.  
 XX PI Van Den Berg MA, Bovenberg RAL, Driessen AJM, Konings WN;  
 XX PI Schuurts TA, Nieboer M, Westerlaken I;  
 XX PR WPI: 2001-291055/30.  
 XX DR N-PSDB; AAF83399, AAF83400.  
 XX PR Enhancing secretion of beta-lactam compounds from a micro-organism by  
 PT enhancing adenosine triphosphate-binding cassette transporter activity,  
 PT useful for producing e.g. penicillin and cephalosporins -  
 XX Claim 8; Page 110-112; 116pp; English.  
 XX CC The invention relates to a method for enhancing the secretion of beta-  
 CC lactam compounds from a micro-organism that comprises enhancing ABC  
 CC (adenosine triphosphate (ATP)-binding cassette) transporter activity of  
 CC the micro-organism. The method is used for enhancing the production and  
 CC secretions of beta-lactam antibiotics like penicillin and cephalosporins.  
 CC Manipulation of the ABC-transport protein system provides a means for  
 CC enhancing beta-lactam secretion. The present sequence represents the  
 CC P. chrysogenum ABC transporter aa10 protein sequence.  
 XX Sequence 1333 AA;  
 SQ  
 Query Match 30.0%; Score 1934.5; DB 22; Length 1333;  
 Best Local Similarity 35.3%; Pred. No. 1.7e-171;  
 Matches 490; Conservative 226; Mismatches 468; Indels 203; Gaps 28;  
 QY 5 SEKPRPNQDSESSHPPEKTEGSI-----DYL 35  
 Db 21 nekpvdrdsaststakehsktpvasdedalyahpehekdlkqldappvnisyf 80  
 QY 36 RIFRYADKYDWTUNVIALICAGSASLPMSIIFGSFTNKNFNNSGDSPEAFKADVD 95  
 Db 81 glyryasridlilivstlcaiaagaalpfiltifgslatafqlmrltipydefydl 140  
 QY 96 HFVLWFVYLFGRFVLTYYSTAATISARTFTLRVFELECTLRQEWHPDKQNGAIA 155  
 Db 141 tnvlyfyigerevtvystvgfytygehvqtqirehyaleallrqnlayfdkigagv 200  
 QY 156 TXVTTNGNRIGTGAELKFTVQALSMTFFSAFVVALASQWKALITMSVIPAFIVTG-- 213  
 Db 201 tritadnllqdgvsckvgitlcvatftvafvayikyaplagictsmvalvilmggg 260  
 QY 214 -----ICTAIDAAQEARITRIYSRAAVLAEEVLSSIRTVAFYAOKKMVEKYDFV 266  
 Db 261 sqllikfgkksiesea-----gaggtvaaeevisirnatatfgtdkikayeshllr 311  
 QY 267 AHQCKKKSPNNGVLFSTEFYCIYAAALAFNKGFRMYONGEVADVGVFTTVVLSVTLAA 326  
 Db 312 aervgmalqmslvmvglmfmnyglfwmgsqlvdgkv-dwggvltlmailigs 370  
 QY 327 TSISMLAPSGSVVYQRRIFGSELFSLIDKPTOLDPLDPSGKQPEGCLGQIEIQNLAFAP 386  
 Db 371 fslgnvspnasaftnavaatkifatlrdslpdsldptsdegildvheghiefrnvkhiyp 430

QY 387 SRPSAOVLRFENLTIPAGKTTALTAVGASGSKSTMVGLLERWYLPSSGRILLDGLGOYN 446  
 Db 431 srpevtmndvslgipagntaltavpsgskstvglyverfylpvggvflgdhdtcln 490  
 QY 447 VKWLSRIKLVQOEPVLPFGTIFQNIANGFMDQRLDPRKOM-ELVQKACKASNGDVFI 505  
 Db 491 lrwlrqsilvsqepvifgttiyqirhgligrfheseeekikellenaakmanahEFI 550  
 QY 506 NELPNGYETEGERAGALSGGORIATARSISDPKILLDDEATSLADPKAEKVVQEA 565  
 Db 551 talpegyetnvqrgfllisgggkqriataramvdpkllldeatsaltdtsegvvaal 610  
 QY 566 NRVSQDRTTLVIAHKLATVKSAGNIATVISOGKIVQGTTHHELIFGCHYAALVRAQ 621  
 Db 611 draaegrttiavahrstikshahnvfvqgsivqegthsltehdgpyfvlveaqrine 670  
 QY 622 -----DLGADEQOEH-EKTLHEKAAR-----EAGGERPALERTHTTATSOAG 662  
 Db 671 ekdadaladededgleemtkshiarvksiasgstcvkdeaeetfgdamhr-gesrksvss 729  
 QY 663 DLEKRRVPVGTGLYSLKILIMFYEQKNLWCFLLSTITVLICANATPGQALLFSRLLT 722  
 Db 730 vilsqktaeggrkshliltlikfigsnkeerwfmaiglcfsilagcgqptqaflyakais 789  
 QY 723 VFEL-----SGHAAQRADEFYILMFFVVALGNLVGYFTIGTMCNVISQVVTHTHYOAMFORV 779  
 Db 790 alsipktqydklrsdanfwslmfivqviitfsvhglafafaserlirkargnafrvm 849  
 QY 780 LDQDIELLDIPQISGALTSQSLPTQLOEL--ISANFLIYIVVQHLRBOCSTTSIWM 837  
 Db 850 lrqdlfnffreenstgaltfsls---teckhlaglsqgtl----- 886  
 QY 838 ETGPGCVCTCTTP-----AFGWLPQNSRDEARSKLGLCKR----- 876  
 Db 887 -----gtlmtsttlasiavialafgw-----klalvcmvsiipillgcgyr 928  
 QY 877 -----CWACKRSS-----YRDPDLIPDSRRPCSPVLGHVE- 908  
 Db 929 fymlaafqarskaayegsasyacealsairtvaeltretvwsf-----yhgqler 979  
 QY 909 QGLAKIIOSFWFCRFGFHLQSMEFLAIALGFCIAVDNWLQVSTQLN-----FI 959  
 Db 980 qgrtslilvfkssliya--asqalvfcvalgf-----w--yggllighbeydvfrfv 1029  
 QY 960 SSSWAFCLPVQAAQIYAVSTFTKARSANVILWLTLPKPTIRE-TEENKKKGPVGGCP 1018  
 Db 1030 cfs-eilfagagagtvsfspdmgkaknaaaeflrlferrptidtwseegenldhceg-t 1087  
 QY 1019 VDLEDIEFRYORDSARVLRGVSMTEPQQFVAYVVGASCGKSTLIALSERFYDPTSGRI 1078  
 Db 1088 iefkdvhyfyrtpreqpvirgintvbkpqqyalvpgsgcksttiallerfydalsggv 1147  
 QY 1079 SPAHENTAESPLRYGHMSLQOEBTLYOGSVRENNVTLAL--EAEELSEELCOGRLPARP 1136  
 Db 1148 yidkniadlnvnsyrschialvsqetlyvgtkikenillgspnqdpqteelvgvckdani 1207  
 QY 1137 MWLISLQKALKRLAAORGWQFSGGQQRQRIATIRALIRNPKLLLLDEATSLDTSER 1196  
 Db 1208 ydfims--lpegfntvvgsgkmlsggqkqrvaiaarallirnpkvlildeatsaldsesek 1265  
 QY 1197 LVQAALDEASTNTTIAVAHRLSTIRNVDFIEVFANGRIAEFTGHAEQLRLGRYEMCL 1256  
 Db 1266 vvaaladaargrttiavahrstlqkadilyyfdggkivesgthtellrnkrgyfeivn 1325  
 QY 1257 AQSLDQA 1263  
 Db 1326 lqslgra 1332

RESULT 7

AAW62871

ID AAW62871 standard; Protein; 1408 AA.

XX AAW62871: 27.7%; Score 1789.5; DB 19; Length 1408;  
AC Best Local Similarity 34.4%; Pred. No. 7.9e-158; Indels 203; Gaps 35;  
DT Matches 478; Conservative 214; Mismatches 493;  
XX 29-SEP-1998 (first entry)  
XX Multiple drug resistance protein called CnemDR1.  
XX Multiple drug resistance protein; CnemDR1; Identification;  
KW antifungal compound; resistance.  
XX  
XX Cryptococcus neoformans.  
XX US5786463-A.  
XX 28-JUL-1998.  
XX 08-MAR-1996; 96US-0612521.  
XX 08-MAR-1996; 96US-0612521.  
XX (ELIL ) LILLY & CO ELI.  
XX Peery RB, Skatrud PL, Thornevell SJ;  
XX WPI: 1998-436610/37;  
XX N-PSDB: AAV42347, AAV42348.  
XX Dna encoding multiple drug resistance protein of Cryptococcus  
PT neoformans - also related vectors and transformed host cells, and  
PT method for expression of the protein  
XX Claim 1; Columns 17-26; 19pp; English.  
XX  
XX The present sequence represents a multiple drug resistance protein of  
CC Cryptococcus neoformans, designated CnemDR. Host cells transformed with  
CC the DNA sequence are useful for identification of potential antifungal  
CC compounds, and for enhancing the activity of proven antifungal compounds  
CC to previously resistant species.  
XX Sequence 1408 AA;  
XX  
Query Match 27.7%; Score 1789.5; DB 19; Length 1408;  
Best Local Similarity 34.4%; Pred. No. 7.9e-158; Indels 203; Gaps 35;  
Matches 478; Conservative 214; Mismatches 493;  
QY 6 EKPRPDGSESSHP-----PPEKTEGSI---SDYLIRIFRYA 41  
DB 92 ekplsn---afskshpykkskdfllksrkkkeeeeknkekeasvlpvpsffalrfra 148  
QY 42 DKYDWTUNVIALICAGSGASLPLMSIIFGSEFTNKNNY-----NSGDGSPGA----- 89  
DB 149 apileiamvlgvlavaagscplmlflgrltsfnyavtanqlsqggltpetsaaalq 208  
QY 90 -----FKADVDFVLFVIFIGKFLVTVYSTAAITISAIRTRRLRRVFLCETLRQEVW 144  
DB 209 aakddlktqghnalylmaigmflatlwlmfwnvtgelnskrirerylaavlrqeta 268  
QY 145 HFQKQSGALATXVTNGNRIOGTGIAEKLVFTVQALSMFSPFVVALASQWKLALITMSV 204  
DB 269 yfdldgagavatigtdchivqegtskvalvfqyagtfvgfvlavfvrspriagelvali 328  
QY 205 IPAIFLVTGICIAIDAAQAEARITRIVSRAAVLAEEVLSSTIRTVHAYAQKKMVKYDVFL 264  
DB 329 lpvmlcggkmmkamakgygtaaldhakagslaeevlgslrtvqafgkeklgdkfadhli 388  
QY 265 QQAHEGKKKSPNNGVLFSTPEYFCIYAALAFWKGFRMYQNGEADVADGVFTVVLVSPTL 324  
DB 389 eqskivgrkgsifegfslgmffviyaalaffvgilvnsqg-adsqivinvfmslii 447  
QY 325 AATISMLAPSGSVVQRIRFGSELSITDKPTQLDPLPSGKQPEGLGQIEIQMLAPA 384  
DB 448 gsfmamlapelaavtkargaaaklfatidrvpaldsaseegfkpdlrgelgsfenvkfh 507

QY 385 YPSRPSAOVLRDNLTIPAGKTATLVAGSGSGKSTMVGLLERWYLPSSGRILLDGLGLEQ 444  
DB 508 ypsrpsipilkgtttfeagktfalvgasgsgkstvslierfydpsvgvklrdgrfirs 567  
QY 445 YNVKWLRSRIRLVQOEPVLFRCITFONTANGFMDSQ-RDLPREKOMELVQKACKASNDV 503  
DB 568 llnwlrqqglvsqepdlfgttvrgnvengligaryenasleekfelvkkacvanahn 627  
QY 504 FINEPNGYETEVGERAGALSGQRRTAIARSIIISDPKILLDBATSALDPKAEKVQOE 563  
DB 628 fmkpipygydtmvgermilsggqkrvaiairaivsdprillldeatsaldtqsegivqd 687  
QY 564 ALNRYSKDRRTLLVIAHKLATVKSAGNIAVISOGKIVEQCTHHELI--EFGCHYAALVRAQ 621  
DB 688 aldksgrttitiahristirdadriyvmgggvelegshndllaneng-pyeqlvnnq 746  
QY 622 DL-----GADEQOEHEKTLHE-----KA 639  
DB 747 klqaeaaealqvddiedpddavfigsspmqekdkqlhravtgrslasiamddiqakr 806  
QY 640 AREAAGERPALERTHTTATSQAGDLEKRVKVPVGTLCYSLKLCILIMFYEQKNLYWCFLLS 699  
DB 807 acevage-----dkip---sfglyarllrmnsadkfiy---ifa 840  
QY 700 TTVVILICAAATPPGQALLPSRLTLVPELSG-----HAAERADPFYILMFVVALG----- 748  
DB 841 flaaicagmvpyslaifgalsdfeiqdpaelhalarsea-----lwyfitalaafvif 896  
QY 749 -NLVGYFTIGTCNVISQVTHRYQAAAFORVLDQIDELLDIPEQIISGALYSQLSALPTQ 807  
DB 897 fgsagfsragwdlmgv-----lrrkkliftatlrhdiewfdeernstgavtsanldqpgk 949  
QY 808 LQELISANFLYIVVGQHRLEBOCSTTSLWMTGPGCC-----VMCTSTPAP-- 853  
DB 950 vggifgptgtv-----gscatli-----ggcligicygplaligtaciipilvs 995  
QY 854 -CWL-----PQNSSRDEARSRKLGLCRKWCACKRSSYRDPDLRIFDSRRPSPVTL 904  
DB 996 ggyirllkvvlkdqrmkklhaasahlaseagavktvasltrek-----dvrrlysealk 1050  
QY 905 GHVEOGLAKITOSFWFGRCFGPHLSQSMEFLAIALGFCIAVDNWL---OVSTQLNFIS 961  
DB 1051 apmklnfrtsiks---qcl-faasggitcfclialivfyiga-lwldakystasfytvin 1104  
QY 962 SWAFCLPVQAAQYLAYSTFTKARSAANYILWLRTLKPTIR-EPEENK---KKGPGVGC 1017  
DB 1105 slvfa-siq-agnvftvdpaskausaasifrsldnepainaesnegkvidhkhvvg-- 1160  
QY 1018 PVDLEDIEFRYRQDSARVLGRVSMTEBPQOPVAVVVGSGCKSTLIALSERFYDPTSGR 1077  
DB 1161 hvriegvhfrptprgvrvtlnltidvpagtyvalvgpgcgskttigmlerfydplagr 1220  
QY 1078 ISFAHENTAEKSPRILYRGHMSLVQOEPVLYOGSVRENTVLAELSEELCOGRUPAR-- 1135  
DB 1221 vtlldgidikealnasyrsqslvsqetlyagtirfnllgankpl-eevtqdeidaack 1279  
QY 1136 -PMLWLSLYQKALKRLAAORGQFSGQQRORIAARALIRNPKLLLLDEATSLDPTQS 1194  
DB 1280 daniydfivspdgtdtevgkgqslsggqkrataraalirnpkvllldeatsaldsqs 1339  
QY 1195 ERLVQAAALDEASTPTTIAVAHRLSTIRNVQVIFVFANGRIAEQTGTHAELORLORYEM 1254  
DB 1340 ekvvqeadkaakgrttitiahrlslslqhsdrivvfyfsegrvnaehgthqellakkggyvel 1399  
QY 1255 CLAQSLDQ 1262  
DB 1400 vqmqnlr 1407

RESULT 8  
AAR99255  
ID AAR99255 standard; Protein: 1307 AA.

```

XX AAR99255;
XX 01-DEC-1996 (first entry)
XX Aspergillus flavus multidrug resistance protein Afl-MDR1.
XX Multiple drug-resistance; MDR; Afl-MDR1; antifungal.
XX Aspergillus flavus.
XX W09626952-A1.
XX 06-SEP-1996.
XX 23-FEB-1996; 96WO-US02665.
XX 27-FEB-1995; 95US-0395246.
XX (ELIL ) LILLY & CO ELI.
XX Peery RB, Skatrud PL;
XX WPT; 1996-412733/41.
XX N-PSDB; AAT35230.
XX New multiple drug resistance gene of Aspergillus flavus - used to
XX identify cpds. that inhibit fungal multiple drug resistance activity
XX Disclosure; Page 22-27; 37pp; English.
XX A new multiple drug resistance (MDR) protein (AAR99255), designated
XX Afl-MDR1, of Aspergillus flavus was identified as the product of
XX a cDNA clone isolated from a mycelial cDNA library of A. flavus.
XX Afl-MDR1 can be synthesised by host cells, partic. Saccharomyces
XX cerevisiae, transformed with vectors that provide for the expression
XX of DNA encoding Afl-MDR1. Transformed host cells can be used to
XX identify cpds. that inhibit fungal MDR activity.
XX Sequence 1307 AA;

Query Match 27.6%; Score 1780; DB 17; Length 1307;
Best Local Similarity 34.3%; Pred. No. 5.4e-157;
Matches 466; Conservative 222; Mismatches 500; Indels 172; Gaps 32;

QY 7 KPRPNQDSESS-----SHPPP-----EKETEGSISDYLRIFRYADKYDWTNLVIALICA 56
Db 10 kplpkspgtsttgthavshaevevldrqlhtpvsqigfgfgyrtrwvdlfgsala 69

QY 57 IGSASLPLMSIIFGSFTNKNFNNSQDGSPEAFKADVDHFVLMFVFLYFGKFLTYVST 116
Db 70 laggaa1pftvlgfritstfqdiathritdydhfheltnkvvyfiylgaaefvalylat 129

QY 117 AATISARTTTRTERRVLECTLRQEWVHFQKQNGAIAATXVTNGNRITOTGTAERLVFT 176
Db 130 vgfylgdhvvqqlrveyfgallrtnlaftdtlgageittricitadnltqdgisekvgl 189

QY 177 VOALSMFFSAFVALASQWLALITMSVIPAFILVTGICIAIDAQAQEARITRIYSRAAYL 236
Db 190 ltgstftvtafiayiknwklalicsallaltlmgcstlmifskkalelyqvggasm 249

QY 237 AEEVLSSRTTHAYQAQKVEKYDVELQQAQHEGKKKSPNGVLFSTEFYCIYAALIA 296
Db 250 aedliidsrtvaaadqatarkyeshlkdaegpgmkskvifaimvgallclmlynlg 309

QY 297 FMKGRMYQNGEADV--GKVTYVLSVTTLAATISISMILAPSGSVVYQRRIFGSELSIID 354
Db 310 fmgsgfriveg-isnikagdvltimmaillgynlgnvapgqalsdavaaasklygtid 368

QY 355 KPTQLDPLDPGKQEGCLGIEIQLNLAFAYPSPRSQVLRDFNLITIPAGKTTALIVGASG 414
Db 369 rqspldalsdgdktlefrvgnlviqnrhyvpsrpevtvahndiscyipagkttafvgpsg 428

```

```

QY 415 SGKTMVGLLERWVLPSSGRILLDGLGELGOYNVVKWLSRLRVQEPVLFRTQIFQNTIAN 474
Db 429 sgkstiislllerfydpvagtimdgndigtlnlrwrqmslvsgqepfrlftaenry 488
QY 475 GFMDQORDLPREKOMEL---VQKACASNDGVFINELPNGYEYEVGERAGALSQGRQRI 531
Db 489 gligr--fekestyeirkrveaaarmanahdfimalpnqydtnei--esfslsggqkqi 544
QY 532 AIARSIISDPKILLDEATSDALDPKAKVVOEALNRVSKDRTTLVTAHKLATVKSAGNTA 591
Db 545 aiaraiikdpkllldeatsldtkscklvgaaldkaskgrttivlahrlstiqkaynii 604
QY 592 VISQKIVEQTHHELIEFGCHVAALVRAODLGADGDEQOEHEKTLHEKAAREAAAGERPALE 651
Db 605 vlangivegqgphelmdrrgiycdmveahei-----kkryaryskrysqlltnls 655
QY 652 RTHTTAT-----SQAGDL-----EKKKVPVOTLGVSLKKGILIM 685
Db 656 pkhnpmtffdkdypgddesdiysilssddasdiglhtgekqr-pvarmelshlmpvkee 714
QY 686 FYEQKNLY-----WCPLLSTITVLICAATF-PGOALLFSRLTLTFE---LSGHAA 731
Db 715 aysfwtkfklflasnfpewpfilllgicassllagigpsqavlfakavstslsppleykpl 774
QY 732 QERADFYILMFVVALNLNVGYFTIGWTCNVISOVTHRYOAMFORVLDQDIELLDIPE 791
Db 775 rhdanfclmflmigivslvlysvqgtlfayssmekmyrarsqafvilhqdisffdgqe 834
QY 792 QISGALTSOLSALPTOLQE-----LISANFL-----IYIVVGQH-----RL 827
Db 835 ntgallatlasagtkeltgsgvtgltilivsnlvlasglvalvlgwkialvcisavpal 894
QY 828 EQCSTTSLMBETPGCGVMTCTSTPAFWLQNSRDEARSKLGLKCRKQWACK-RSSYR 886
Db 895 lmcgfvrvwm-----lerfrrakakayqesases-aceaasair 931
QY 887 DPDLRIFDS-----RPPCSPTVLGHVEQGLAKIIQSFWRGRCFGPHLSOSMEFLA 936
Db 932 tvvsltmetealqsyqaqlrrqlksdiipvksll-----yassqaipffc 978
QY 937 IALGCTAVDNWL-----QVSTTQINFISSMAFCLPVQAAQVLAYSTSTFTKARSA 988
Db 979 malgf-----wyggsllghgelysalfgyvcfseviv--gagaagtvsfshapdmgakha 1030
QY 989 ANYILWRLTLKPTIRE--TEENKKKG-PVGGCP--VDLEDIEFRYRORDSARVLRGVSMT 1043
Db 1031 a-----refrlfssdtmhasrskgvvvtmsmrglvfsvdsvfrypsrleqpilrhlnit 1084
QY 1044 IEPGOFVAYVAGSGCGSTLIALSERFYDPTSGRISPAHENIAEMSPRLYRGHMSLVQOE 1103
Db 1085 lkpgqfvalvgasgsgksttiallerfydpkpgvvyvvgdkniltlemssyrshlaisqe 1144
QY 1104 PTLXGGSVRENTLAEAE--LSEELCOGRLPARPMLWLSSLYQKALKKLAQRCGMQFS 1161
Db 1145 ptlfggtirenlignstphvtdtflvackdaniydfils--lpqgfnitvgnkggmls 1202
QY 1162 GQQRQRIARALIRNPKLLLDDEATSDALTOSELRLVQALDCASTRRTTIAVAHRLSTI 1221
Db 1203 gqdkriaalarilrnpkllldeatsldtsesekvsvvgaaldaaargrttiavahrlist 1262
QY 1222 RNVDVIFVFANGRIAETGTTHAELQRLGRVYEMCLAQSLD 1261
Db 1263 qradliyvldgvevsgsthrcllrkkgryyelvhlqopd 1302

```

```

RESULT 9
AAB81065
ID AAB81065 standard; Protein; 1283 AA.
XX
AC AAB81065;
XX
DT 25-JUN-2001 (first entry)

```

XX Cynomologous monkey P-glycoprotein variant 2.  
 DE Cynomologous monkey; P-glycoprotein; GPG; multidrug transporter; MDR1;  
 KW efflux pump.  
 XX Macaca fascicularis.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 93..95  
 FT /note= "An additional 3 amino acids are present compared  
 FT to PGP variant AAB81064"  
 XX  
 PN WO200123565-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26592.  
 XX  
 PR 28-SEP-1999; 99US-0156921.  
 PR 12-OCT-1999; 99US-0158818.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 PI Stocker PJ, Steimel-Crespi DT, Crespi CL;  
 XX  
 DR WPI: 2001-316136/33.  
 DR N-PSDB; AAF86128.  
 XX  
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 PT cell -  
 XX  
 PS Claim 9; Page 65-68; 84pp; English.  
 XX  
 CC This invention relates to a polynucleotide sequence encoding a  
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention  
 CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents the  
 CC cynomologous monkey P-glycoprotein variant 2. The protein has an  
 CC additional 3 amino acids when compared to PGP variant 1 (AAB81065).  
 XX  
 SQ Sequence 1283 AA;

Query Match 25.7%; Score 1656; DB 22; Length 1283;  
 Best Local Similarity 32.6%; Pred. No. 2.3e-145;  
 Matches 428; Conservative 252; Mismatches 495; Indels 138; Gaps 28;  
 QY 24 EKETEGSILURIFRYADKYDWTNLVIALICAIGSGASLPLMSTFGSEFNK----- 76  
 DB 27 kkerkptsvf-smfrysnwldklymvgvtlaaihagblplmmlvfgdmdtdtfanaqnl 85  
 QY 77 -----FNMYNSG---DGSPA-PKADVDHFLMFVYLFYFGKFTVLTYSTAAITSAIR 125  
 DB 86 galgallfntnssnltcvpmleedmtrvayyyysgigagvlvaaylqvswfswclaagr 145  
 QY 126 TTRTLRRVLECLTQROEVYHFDKQNGAIAATXTVTNGNRIOTGIAEKLVTFTVQALSMFES 185  
 DB 146 qihkirkqffhaimrqelgwfdvhdvgeintrltdvskinegldkigmffqsmatftt 205  
 QY 186 APVVALASQWKALMTMSVIPAFIVLTGICIAIDAAQARITRIYSRAAVIAEEVLSIR 245  
 DB 206 gfivgfrgwkltvilaispvgisaaavwakilssfdkellayakagavaeevlaair 265

QY 246 TVHAFYAQKMKVEKYDVFLOQAHOEGKKSPNNGVLFSTEYECIYAAIALAFWKGRMYO 305  
 DB 266 tviafggqkkelerynkleeakrigikaltanisaigaafiliyasyalaewygttlvl 325  
 QY 306 NGEVADVGKVFVTVSVTLAATISMLAPSGSVYQRRIFGSELSIIDKPTQLDPLDPS 365  
 DB 326 skeys-igqvltvfvsvligafsvgqaspsleafanargaafeifkiidnkpdsiayks 384  
 QY 366 GROPEGCLQOIEIQNALFAYPRPSAQVLRDNLTIPAGTKTALVAGSGSGKSTVWGLLE 425  
 DB 385 ghkpnikgnlefrnvhsygsrkevkiklgnlkvgsgqtvalvngsgcstvtqlmq 444  
 QY 426 RMYLPSGRIILLDGLLEGOYNVKKWRSRLVQOPEPVLPRGTIFONIANGFMPDEORDLPR 485  
 DB 445 rlydptegmvsvdgqdirtnvfrlreilgvsvsgepvifattlaenirygred----- 497  
 QY 486 EKOMELVQKACKASNGDVFINELPMGYETEVEGERAGALSGGORQRIATARSIIISPKILL 545  
 DB 498 -vtmdeiekavkeanaydfimklpqkfdtlvgersqglsggqkqrlaaraivrnpklll 556  
 QY 546 LDEATSAIDPRAEKVVQBAELNRVSKDRITLVIAHKLATVKSAGNIAVTSQGIKIVGQTHH 605  
 DB 557 ideatsalteseeavvqvaldkarkgrttivlahrlstvrnadvlagfdgvgivekgnhd 616  
 QY 606 ELIEFGCHYAALVRAQDLGADDOOHEKTLHEKAAAREAGERPALR-RTHTTATSOAGDL 664  
 DB 617 elmkekglyfklytmqtagnetelelen-----aadeakseldtlemshdsgsalikr 669  
 QY 665 EKRKVPVGTLG-----YSLUKCILIMFYEQKNLYW-CFLLSITTV 703  
 DB 670 strfsvrgsgqdrklstkealdeisppvsvfrlmklnl-----tewpyfvvgvfca 721  
 QY 704 LICATFEGQALLFGRLLTVFELSCHAQER--ADFYILMPFVALGNLVGFTTGWTCN 761  
 DB 722 lingglqpfavifskiiigiftrnddaetkqnsnlfalflivqivafitflgffgfg 781  
 QY 762 VISQVTHRYQAMFORVLDQDIELDITPEQISGALTQSLSALPTQLOELISANFLIIV 821  
 DB 782 kageiltkrlrymvfrsmrlrqdvswfdqpkntgtalttrlandaaqvgkgaigrilai-- 838  
 QY 822 VQHRLEQCSVTSIWMETGPGCVMTSTPAFGW-----LPO 858  
 DB 839 -----ltqnlaniigtglll---sllygwtlltllaiplaiagvvenmklag 884  
 QY 859 NSSRDEARSLKLCRCWACKSSYRDPDLRPFDSRRPCSPVLGHV-EOGLA----- 912  
 DB 885 qalddkkelegagkiatea---ienftvsvltqeqk-----fehmydgsiqvpyrn 933  
 QY 913 KIIOSFWFGRCFGFHLQSOMBEFLATALGFCIAVDNMLQVSTQNLNFISSSWAFCLPV--- 969  
 DB 934 slrkahifgicfsf--tqammyfsya--gcfrfgayl--vahalsmfedvllvfaavfga 988  
 QY 970 QAAAOYLAYSSTPYKARSAANYILMLRTLKPTIRE-TENKKKGPGVGGCPVDLEDIEFRY 1028  
 DB 989 mavqgvssfapdyakakavsaahililektplidsystegikpntleg-nvtfnevvfny 1047  
 QY 1029 RORDSARVLGVSMTEPGQFVAVVGASCGCKSTLIALSERFYDPTSGRISFAHNIAM 1088  
 DB 1048 ptridipvlqlsltevkkgqlalvgsqcgkstvwqllerfydplagkvlldgkelqkl 1107  
 QY 1089 SPRLYRGHMSLVQOEPPTYLQGSVRENVTALAEAL--SEELCQGRLPARMLWLISLSYQ 1146  
 DB 1108 nvqlwrlahlgivsqeplifdcisentaaygdnsvrsveeivraakeanihaf1-eslpn 1166  
 QY 1147 KALRLAAQGMQFSGGQRIATARALIRNPKLLLDDEATSAIDTQSERLVQOALDEAS 1206  
 DB 1167 kystrv-gdkgtqlsggqkqrlaaraivrphillideatsaltesekvvsqealdkar 1225  
 QY 1207 TSRTTIAVAHRLSTIRNVDFVFPFANGRIATETGTTHAEQLRGRYEMCLAQ 1259  
 DB 1226 egrtcivlahrlstiqnadlllvfnggrvkehgtthqqlaqkgyfismvsvga 1278



```
RESULT 10
AAB81064
ID AAB81064 standard; Protein; 1280 AA.
XX
AC AAB81064;
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein variant 1.
XX
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW efflux pump.
XX
OS Macaca fascicularis.
XX
PN WO200123565-A1.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26592.
XX
PR 28-SEP-1999; 99US-0156921.
XX
PS 12-OCT-1999; 99US-0156818.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
PI WPI; 2001-316136/33.
PI N-PSDB; AAF86127.
XX
DR
XX
PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT (PGP) and homologous PGP polypeptides are useful for predicting
PT bioavailability of compound and increasing PGP transporter activity in
PT cell -
XX
PS Claim 9; Page 57-59; 84pp; English.
XX
CC This invention relates to a polynucleotide sequence encoding a
CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC exporting small molecules across the cell membrane. The invention
CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC and protein, and also that of an allelic variant. The PGP polynucleotide
CC sequence is useful for increasing PGP transporter activity in a cell.
CC Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC activity in a mammalian cell. They may also be used for increasing the
CC bioavailability of a drug. The present sequence represents the
CC cynomologous monkey P-glycoprotein.
XX
SQ Sequence 1280 AA;
Query Match 25.6%; Score 1654.5; DB 22; Length 1280;
Best Local Similarity 32.5%; Pseq. No. 3-2e-145;
Matches 426; Conservative 253; Mismatches 496; Indels 135; Gaps 27;
QY 24 EKETGGSIDYLRIFRYADKYDWTNLNVALICATGSGASPLMSIISGSPTNKF----- 77
DB 27 kkerkptsvsf-smfrysnwldkiymvvgvtlaaibhagiplmmlvfgdmdtdfanagnl 85
QY 78 -----NNVNSGDSPEA-FKADVDHFVFLWFLFGRFVLTYYVSTAAITISAIRTR 128
DB 66 gdlgalltssnntdtpvmnleedmtryayvysgigavlvaayiqvswfclaaagqlh 145
QY 129 TLRVRVLECTLHQEVWHFDKSGNAIATYVTTNGNRITOTGIAKLVFTVQALSMFYSAFV 188
DB 146 kirkqdfhaimrqeigwfdvhdvgelntrltddvsknegldkigmffgsmatfftgfi 205
QY 189 VALASQWKLALITMSVIFAFLVTCICIAIDAAQEAIRITRIYSRAAVLAEEVLSIRTVH 248
DB 206 vgftrgwktlvilaivspvlgsaavwaklissftdkellayakagavaeevlaartvl 265
```







```
Db 882 qalkdkkelegagkiatea-----ienfrtvsvltqeqkfehmyaqalqvp 927
QY 917 -----SFWGRCGFGHLSQSMFLAIA-----LGFCIAVDNMLQVSTTQLNFISSWAF 965
Db 928 yrnslrkahiifitfsf--tqammyfayagcfrfgyalviahkmsfedvllvfs-----av 981
QY 966 CLPVOAAQYLAYSTSFTKARSAANYILWLRLTKPTT--RETEENKKKGPGVGCPCVDLE- 1022
Db 982 vfgamavqvsfapdyakakisaahimiektplidsyste-----gimpntleg 1033
QY 1023 -----DIEFRYQDSARVLGVSMTIEPGQFVAYVCASCGKSTLIALSERYDPTSGR 1077
Db 1034 nvtfgevinyptpdpipvlvglsievkkkgqctialvsgsgckstvvqlrlfydplagk 1093
QY 1078 ISFAHENIAEMSPRLYRGHLSVQEQPTLYOGSVRENVTLALAEAL--SPELCQGRLPAR 1135
Db 1094 vildgkeikrinvqlrahlglvsgqepillfocslaenaygdhsvvsggeivraakean 1153
QY 1136 PMLWTLSSLYOKALKRLAAQRMQFSGGORIATARALRNPKLLLLDEATSAIDTQSR 1195
Db 1154 thafieslpnkystk--vgdkgqtqsggqkqialaralvqphillldeatsaldese 1211
QY 1196 RLWQALDEASTRTTIAVAHRLSTIRNDVDFVFANGRIATCTHAEQLRLGRYYEMC 1255
Db 1212 kvvqealdkaregrtciviahrlstignadlivvfgqrvkehgthqqlagkgyfsmv 1271
QY 1256 LAQS 1259
Db 1272 svqa 1275
```

## RESULT 14

AAR04868  
ID AAR04868 standard; protein: 1280 AA.

XX AAR04868;

XX AAR04868;

DT 01-OCT-1990 (first entry)

DE Protein encoded by Multidrug Resistance A1 gene.

KW Multidrug Resistance A1 gene; drug resistance of cancer cells; ss.

OS Homo sapiens.

PN JP02100680-A.

XX 12-APR-1990.

XX 05-OCT-1988; 88JP-0251475.

XX 05-OCT-1988; 88JO-0251475.

XX (SUNR) SUNTORY LTD.

XX WPI; 1990-159707/21.

DR N-PSDB; AAQ04522.

XX MDR related gene derived from human normal cells -

PT codes specific amino acid sequence, used for diagnosing drug

XX resistance of cancer cells

XX Disclosure; ; p; Japanese.

XX The gene that encodes this protein is useful for diagnosis of drug

CC resistance of cancer cells.

XX Sequence 1280 AA;

Query Match 25.5%; Score 1647.5; DB 11; Length 1280;

Best Local Similarity 32.3%; Pred. No. 1.4e-144;

```
Matches 428; Conservative 242; Mismatches 491; Indels 163; Gaps 26;
QY 24 EXTEGEGISDYLRFRVADKYDWTNLNVIALICAIGSASLPLMSIIFGSGFTKFNFN---- 79
Db 27 kkekptsvf-smfrysnwldklymvvglaailhgaglpmlmlvfgemtdifanagnl 85
QY 80 -----YNSGDGSPEAF-----KADVDHFWLFWFYLFIGKFVLYVYSTAAITSAIRTR 128
Db 86 edlmsntrnsdindtgffmnlleedmtryayysgigagvlvaayiqvfwclagrqih 145
QY 129 TLRVFELECTLRQEWHEFDKQSNCAIATXVTITNGNRIOTGIAEKLFTVTVQALSMFSAFV 188
Db 146 kirqkfthamrqeigwfbvhdvgeintrltdvskinegidkgimfqsmafttfgfi 205
QY 189 VALASQKWLALITMSVTPAIFLVTGICIAIDAAQEARITRIYRAAVLAEVLSSTRTVH 248
Db 206 vgftrgwktlivilaispvlgsaavwakllsftdeellayakagavaeevlaalrtvl 265
QY 249 AFYAQKKMVEKYDYFLQAOHQEGKKSPNNGVLFSTFEYFCIYAAIALATFWKGRMYQNGE 308
Db 266 afgqkkeleynksleearigikkaitanisaigaaflliyasyalatwygttlvsge 325
QY 309 VADGVKFTVVLSTLAATSISMLAPSGSVYORRTFSGSELSIIDKPTQLDPLDPSGKQ 368
Db 326 ys-igqvlvtvffsligafsvqaspsieafanagaayefiklndkpsidsysksghk 384
QY 369 PEGCLGOIEIQNLAFAPYRSPSAQVLRDFNLITIPACKTTALVGASGSGKSTWVGLLEWY 428
Db 385 pdnikgnlefrnhfsyprskvrlkglnkvsgqtvalvgnsqcgkttvqmqrlly 444
QY 429 LPSSGRILLDGLGQYNVWKLRSIRLVQOEPVLRGTIFONTIANGFMDQBDLPREK- 487
Db 445 dptegmvsdvdqdrtrinvrlreilgvvsgqepvlattiaenirvg-----renv 495
QY 488 QMELVQKACKASNGDVFINELPNGYETEVEGFRAGALSGQORAIARSIIDSPKILLD 547
Db 496 tmdeiekavkeanaydfimklphkfdtlvgergaqisggqkqiaiaralvnrpkllld 555
QY 548 EATSALDPKAEKVVOEALNRVSKDRRTTLVIAHKLATVKSAGNTAVISOGKIVPOGPHHEL 607
Db 556 eatsaldeseavvqvaldkarkgrttiviahrlstvrnadvlagfdgvlvivekgndel 615
QY 608 IEFCHYAALVRAQDLGADEQOEHEKTLHFKAAAREAGERPALERT----- 653
Db 616 mkekgyfklvtmqtagnevelen-----aadeskseidalemsndrsrlkrst 668
QY 654 -HTTATSOAGD-----LEKRVKVPVGLGYSLKCLIMFYEQKNLYW-CPLLSTIV 703
Db 669 rrsrvsgaqdrkistkealdesippvsf--wrimknl-----tempyfvvgvfca 718
QY 704 LTCATFPQOALFSLRLTVFEL--SGHAAQERADFYILMFFVVALGNLVGYFTIGWTCN 761
Db 719 lingglqpafalifskiigvfriddpctrqnsnlslflalgiisfliflqgftfg 778
QY 762 VISQVVTHTYQAMFORVLDODIELLDIPQISGALTSOLSALPTOLOELISANFLIYIV 821
Db 779 kageiftrfymvfrsmirgqdvswfdpdknttgallctrilandaqaqvgkagisrlav--- 835
QY 822 VGOHREOCSTTSLWMETPGGCVMTCTSTPAFGW-----LPO 858
Db 836 -----itqnlantgtgiii-----sfiywgqltlllilaivpiaiaagvvekmisg 881
QY 859 NSSRDEARSRLGLKCRKWCACKRSSYRDPDRILFDSRRPCSPVLGHVBOGLAKIIQ-- 916
Db 882 qalkdkkelegagkiatea-----ienfrtvsvltqeqkfehmyaqalqvp 927
QY 917 -----SFWGRCGFGHLSQSMFLAIA-----LGFCIAVDNMLQVSTTQLNFISSWAF 965
Db 928 yrnslrkahiifitfsf--tqammyfayagcfrfgyalviahkmsfedvllvfs-----av 981
QY 966 CLPVOAAQYLAYSTSFTKARSAANYILWLRLTKPTT--RETEENKKKGPGVGCPCVDLE- 1022
Db 982 vfgamavqvsfapdyakakisaahimiektplidsyste-----gimpntleg 1033
```

QY 1023 -----DIEFRYRQDSARVLKGVSWTIEPGQFVAVVGASGCKSTLTIALSERFYDPTSGR 1077  
 Db 1034 nvtfgvfvnyptrdpvlgdlslevkkgqtlalvsgsgckstvvqlleryfdplagk 1093  
 QY 1078 ISFAHENAENSPRLRYGHMSLVQCEPPLYOGYSVRENTVLALEAEL--SEELCQGRLPAR 1135  
 Db 1094 vllgdgkeikrlnvqwlrahlgivsegeplfdcsiaenlaygdnsrvvsgeevivraakean 1153  
 QY 1136 PMLWTLSLYQKALKRLAAQGMQFSGGQRORIARALINPRLKLLLDENATSLDPTQSE 1195  
 Db 1154 lhaflieslpnkystk--vvgdkgtqlsggqkqriataraalvqrphlllddeatsaldtесе 1211  
 QY 1196 RUVAALDEASTRTTIAVAHRLSTIRNVDFVIFANGRIAETGTHAELQRLGRYVEMC 1255  
 Db 1212 kvvgealdkaregrtciviahrlstiqnadilvvfqngrvkehgthqllaqkgyfismv 1271  
 QY 1256 LAQS 1259  
 Db 1272 svqa 1275

RESULT 15  
 AAB81068  
 ID AAB81068 standard; Protein: 1280 AA.  
 XX AAB81068;  
 AC AAB81068;  
 DT 25-JUN-2001 (first entry)  
 XX Dog P-glycoprotein SEQ ID 7.  
 DE Cynomologous monkey; P-glycoprotein; PGP: multidrug transporter; MDRL;  
 XX efflux pump; dog.  
 KW Canis familiaris.  
 XX WO200123565-A1.  
 XX 05-APR-2001.  
 XX 28-SEP-2000; 2000WO-US26592.  
 XX 28-SEP-1999; 99US-0156921.  
 PR 12-OCT-1999; 99US-0158818.  
 XX (GENT-) GENTEST CORP.  
 PA Stocker PJ, Stelmel-Crespi DT, Crespi CL;  
 PI WPI; 2001-316136/33.  
 DR Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 PT cell -  
 PS Claim 8; Page 73-76; 84pp; English.  
 XX This invention relates to a polynucleotide sequence encoding a  
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDRL is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention  
 CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents  
 CC P-glycoprotein from a dog.  
 XX Sequence 1280 AA;

Query Match 25.5%; Score 1643.5; DB 22; Length 1280;  
 Best Local Similarity 32.2%; Pred. No. 3.4e-144;  
 Matches 423; Conservative 246; Mismatches 504; Indels 141; Gaps 26;

QY 24 EKTEGSIISDYLRIFYADKYDWTNLAVIALICAGISGASLPLMSIIFGFTNKNFNNSG 83  
 Db 26 kkekktvstf-amfysnwlrdlymlvgtmaalingaalplmmlvfgmtdefanagis 84  
 QY 84 DGSP-----EAPKADVDHFV-----LWFVYLF-----IGKPVLYVSTAATITSAIRT 126  
 Db 85 rnkxfvlinesitnntqthfinhleemtyayysgigagvivaayilqsvfswcilaagrq 144  
 QY 127 TRTLRRVFELECTROEWMHFQKOSNGAIATVTTNGNRITQTGAELKLVTVQALSFFSA 186  
 Db 145 ilkirqfhaqrqelgwfdvhdvgeIntrldtdvskinegidkvgnffqsiatffig 204  
 QY 187 FVVALASQWKLALITMSVIPAFPLVGTICIAIDAAQEARITRIYSAAVLAERVLSSIRT 246  
 Db 205 flvgftpgwkltilvialspvlglsaalwakllissfdkellayakagavaeevlaairt 264  
 QY 247 VHAFYAQKMKVEKYDVFLOAHQEGKKKSPNNGVLESTEFYCIYAAIALAFMKRFMYQN 306  
 Db 265 vlsfvggkkelerynklecakrigikkaitanisigaafliliyasyafafwygtsilvs 324  
 QY 307 GEVADYKVTYVLSVTLAATISMLAPSGSVVYQRIFFGSELSFIIDKPTOLDPLDPSG 366  
 Db 325 seyt-lgqvltvffsvlgaftsfgqaspslea fanargaayefkldhkpsidsyeksq 383  
 QY 367 KQPEGGLQIEIONLAPAYSPRSQAQLRDNLTIPAGKTALVAGSAGSKSTMVGLLER 426  
 Db 384 hkpdnkgnlefkvnvhsypsrkevklkglnlkvsgqgtvalvgnsgcgkstvtqlmqr 443  
 QY 427 WYLPSSGRILLDGLLEGQYVNVKWLRSRIRLVQOEPVLFRTGTFONTANGFMDRDLPRE 486  
 Db 444 lydptdmvcidgqdtirtinrvhlreitgvvsgeqelfattiaenlryg-----re 494  
 QY 487 K-OMELVQKACKASNGDVFINELPNGYETEVGERAGALSOGGQRIARISIDPKILL 545  
 Db 495 nvmdeliekavkeansydfimklpnkfdclivgergarls99qkqrlaiaiairvnpnkil 554  
 QY 546 LDEATSALDPKAEKVVOEALNRYSKDRTTLVIAHKLATVKSGNIAVISOQKIVEOCTHR 605  
 Db 555 ldeatsaldtesaavvqvaldkarkgtrtlviahristvnrnadvlagddgviyvekgnd 614  
 QY 606 ELIEFGCHYAAALVRAQDLGADEQOEHEKTLHEKAAAREAAAGERPALE-----RT 653  
 Db 615 elmkekglyfkltmdtrgnelelen-----atgeskesadalemspkdagsliktr 667  
 QY 654 HTTATSQAGDLEKRVKVPVGT---LGYSLALKILINFEYEQNLYW-CFLSTITVLICAAT 709  
 Db 668 strsihapqgqdrk--lgtkedlnenvpavsfwrllkinsteowyfvvgifcailinggl 725  
 QY 710 FPCQALLFRLTLVTF--ELSGHAAQERADFYILMFVVALGNLVGYFTTIGTWCNVLSOVV 767  
 Db 726 qpafsilfrilgtrfdrdpetkrqnsnmfvilvlgisfittflfggftfgkegell 785  
 QY 768 THRYQAMFORVLDQDIELDIPQISGALTSQLSALPTQLQELISANFLIYIVVGOHRL 827  
 Db 786 tkrlrymvfrsmrlrqdvswfdpdkntgtalttrlandaagkgaigrlav----- 836  
 QY 828 EOCSTTSLWMETGPGGCVMCTSTPAESW-----LPQNSRDE 864  
 Db 837 -----lqnlantlgtglli---sllygwltillllaivlialagvvenkmgsgqalkk 888  
 QY 865 ARSRKIKLCRKQACKWACKRSYRDPDRILIFDSRRPCSPTVLGHVQGLAKI10-----SPW 919  
 Db 889 kelegagkates-----ienfrctvvaltrekqfeymraqslqpyvprnsr 934  
 QY 920 FGRCFG--FHLQSMEFLAIA---LGFCIAVDNWNQVSTTQLNFISSSWAFCLPVOAAA 973  
 Db 935 kahifgvysitqammyfsyagcfrfgaylvaneftmfdqvlvifs-----blvfgamavq 990

```

QY 974 QYLAYSTSTKARSAANYILWLRLTKPTIRETEENKKGPVGGCPVDLE-----DIEPR 1027
Db 991 qvssfapdyakakvsaahvimiiekepli-----dsysphglxptlegnvtfnvfn 1044
QY 1028 YRORDSARVLRGVSMTEPCQFVAVYGASGCGKSTLIALSERFYDPTSGRISFAHENIAE 1087
Db 1045 yptrpdipvlogislevkkgqtlalvgsgogkstvvqllerfydplagsvldgkefkx 1104
QY 1088 MSPRLYRGHMSLVQOEPTLYQGSVRENTVLALEAEL--SEELCQGRLPARPMLWILSSLY 1145
Db 1105 Invqwlrahlgivsqepilfdcsiaenlaygdnsrvvshesimqakean-ihhfietlp 1163
QY 1146 OKALKRLAAORGWQFGGQFORIAIARALIRNPKLLLDDEATSDTQSERLVQAALDEA 1205
Db 1164 ekyntv-gdkgtqlsggkqriataralvqphillideatsalidesekvvqealdka 1222
QY 1206 STSRTTIAVAHRLSTIRNVDFVVFANGRIATGTGHAELQRLRGYYEMCLAOS 1259
Db 1223 regrtclviahristiqnadlivvqngkvkehgtqgllaqkglyfsmisvda 1276

```

Search completed: April 27, 2002, 07:45:57  
Job time: 8138 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 20:50:24 ; Search time 5001.19 Seconds  
(without alignments)  
10233.655 Million cell updates/sec

Title: US-09-882-694A-10

Perfect score: 3792

Sequence: 1 atggcagatgaatcgagaa.....aatcttagaccaagcatga 3792

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthm:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estto:\*
- 8: em\_btc:\*
- 9: gb\_estl:\*
- 10: gb\_estl2:\*
- 11: gb\_htc:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	169.4	4.5	687	10	BG646725	BG646725	EST508344
2	168	4.4	897	12	A2541090	A2541090	ENTDS677R
3	167.4	4.4	886	12	BH139685	BH139685	ENTNA477R
4	167	4.4	2676	11	AK014319	AK014319	Mus muscu
5	164	4.3	823	12	A2532602	A2532602	ENTCR50TR
6	162.4	4.3	947	12	A2683753	A2683753	ENTIL96TF
7	161.4	4.3	871	12	A2682350	A2682350	ENTKBI6TF
8	161.4	4.3	919	12	A2690701	A2690701	ENTIY68TR
9	160	4.2	815	10	BI409057	BI409057	602961127
10	160	4.2	834	12	A2548312	A2548312	ENTFG07TR
11	151.6	4.0	906	9	AL578780	AL578780	AL578780
12	150.4	4.0	667	10	BI652986	BI652986	603299956
13	150.4	4.0	899	12	BH154857	BH154857	ENTRO47TF
14	150	4.0	921	12	A2687628	A2687628	ENTIR96TR
15	149.4	3.9	842	12	AZ671925	AZ671925	ENTMN93TR
16	146.2	3.9	672	10	BI176865	BI176865	1031013B0
17	145.6	3.8	832	10	BI904764	BI904764	603168695

18	143.6	3.8	780	10	BG587938	BG587938	EST489713
19	143.2	3.8	943	9	AL520322	AL520322	AL520322
20	142.2	3.7	787	10	BG584063	BG584063	EST485823
21	142	3.7	937	12	BH150760	BH150760	ENTOG42TR
22	140.6	3.7	886	12	AZ540627	AZ540627	ENTGO18TF
23	139.8	3.7	512	9	AV834442	AV834442	AV834442
24	139.4	3.7	1813	11	U66688	U66688	Homo sapien
25	138.8	3.7	635	10	BI176519	BI176519	1031010G1
26	138.2	3.6	853	12	AZ679807	AZ679807	ENTHI67TR
27	137.8	3.6	840	12	BH532776	BH532776	BOHIG29TF
28	137	3.6	691	10	BJ144257	BJ144257	BJ144257
29	136.6	3.6	891	12	AZ682250	AZ682250	ENTKT68TR
30	136.4	3.6	710	12	AQ162582	AQ162582	mqxb0013F
31	134.8	3.6	897	12	AZ550221	AZ550221	ENTGF70TF
32	134.4	3.5	932	12	AZ670821	AZ670821	ENTJN69TF
33	133.6	3.5	931	10	BI415518	BI415518	602969268
34	133.6	3.5	941	10	BI408169	BI408169	602918850
35	133.4	3.5	585	10	BG355606	BG355606	EMI_19_H0
36	133.2	3.5	815	10	BE416931	BE416931	MUG014_G0
37	132.8	3.5	815	10	BE416932	BE416932	MUG014_G0
38	132.8	3.5	880	12	AZ687805	AZ687805	ENTLU52TF
39	132	3.5	605	10	BG694414	BG694414	NTSC_iv01
40	132	3.5	685	10	BJ172661	BJ172661	BJ172661
41	132	3.5	899	12	AZ548799	AZ548799	ENTDK47TR
42	131.8	3.5	750	10	BG585786	BG585786	EST487550
43	131.6	3.5	886	12	AZ668396	AZ668396	ENTHM61TF
44	131.4	3.5	926	12	AZ685308	AZ685308	ENTMS44TR
45	131.2	3.5	621	10	BI264073	BI264073	NF093C09P

ALIGNMENTS

RESULT 1  
BG646725  
LOCUS EST508344 HOGA Medicago truncatula cDNA clone PHOGA-9L3 5' end, linear EST 24-APR-2001  
DEFINITION mRNA sequence.  
ACCESSION BG646725.1 GI:13781837  
VERSION BG646725  
KEYWORDS EST,  
SOURCE barrel medic.  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
REFERENCE 1 (bases 1 to 687)  
AUTHORS Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.  
TITLE ESTs from roots of Medicago truncatula treated with oligogalacturonides of DP 6-20  
JOURNAL Unpublished (2001)  
COMMENT Contact: Michael G. Hahn  
Complex Carbohydrate Research Center  
University of Georgia  
220 Riverbend Road, Athens, GA 30602-4712, USA  
Tel: 706-542-4457  
Fax: 706-542-4412  
Email: hahn@ccrc.uga.edu  
G390331e TIGR sequence name: MTMBQ62TK More information is available at: www.medicago.org  
Seq primer: Skmcd (CTA gAA CTA gtg gAT CC).  
Location/Qualifiers  
1. .687  
/organism="Medicago truncatula"  
/cultivar="A17"  
/db\_xref="taxon:3880"  
/clone="PHOGA-9L3"  
/clone\_lib="HOGA"  
/tissue\_type="3 day old seedling roots"  
/dev\_stage="24 hours after treatment in the dark at 26 C with 0.5 mg/ml oligogalacturonides (DP 6-20) in the

FEATURES  
source



QY 3412 ctttggatttgcctctctttaccagaagcgtttgaaacgccttgcgcctcaacgaggg 3471  
 Db 464 CACATGACTTTATTTTCAGCAATGCCAGAGGATATACAC-AATGGTAGTGATAGAGT 522  
 QY 3472 atgcagtctccggggcgaacagcgaggtatgcgcctatgcgaagcattgattcgaaat 3531  
 Db 523 GCACAAATTTTCAGGAGGACAAAACAAAGAAATTCGCTATTGCCACGCTGATGATTAAGAAAT 582  
 QY 3532 ccaaaagcttgcctacttgacgaagcgacgtcagcctcgcacgcgaatcggaaagctctg 3591  
 Db 583 CCAAAAGTGTTATTAACGATGAAGCTACATCAGCAGCTTGATTCAGAAAAGTCAAAAGATT 642  
 QY 3592 gttaaacgtccctcgtatgaggaatccacagcgccgaacgaacgaatagcagtgccgacgga 3651  
 Db 643 GTACAAAGATGACCTGTGACAAAGCGCAAAAGGAAGAACACAACTGTAATTCACATAGA 702  
 QY 3652 ctttccactatcggaattgattgtttttgtttgttgcgaacggggaatccgcga 3711  
 Db 703 TTATCAACTATTCAAAATGCAGATCAAAATATGTTATTATGAGAGGAAGAAATTCAGAA 762  
 QY 3712 acgggcactcagcggaactcaacagcactgaaggaagatatattacgagatgtgttggca 3771  
 Db 763 ACAGGAACACATCAAGAGTTATTAGATTTCGAAGGATTTTATTATACACTTCTGTATGCA 822  
 QY 3772 caatct 3777  
 Db 823 CAATTT 828

RESULT 3  
 BHI39685  
 LOCUS ENTNA4/7R Entamoeba histolytica 886 bp DNA linear GSS 07-AUG-2001  
 DEFINITION genomic, DNA sequence.  
 ACCESSION BHI39685  
 VERSION BHI39685  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica.  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 886)  
 Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.  
 Determination of clone end sequences from Entamoeba histolytica  
 HMI:IMSS sheared DNA library (2001)  
 Unpublished (2001)  
 CONTACT: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: b.loftus@nih.gov  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
 DNA library  
 Seq primer: M13-Reverse  
 Class: shotgun  
 High quality sequence start: 18  
 High quality sequence stop: 851.  
 Location/Qualifiers  
 1..886  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: pHOSt1; Site 1: Bst I; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica  
 using a method described by Clark and Diamond (Clark,  
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
 method for isolate identification. Exp. Parasitol.  
 77:450.). The DNA was mechanically sheared to give a  
 tight size distribution (~2 kb). The v + i method used for  
 the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for  
 whole genome shotgun sequencing projects. In Genome  
 Sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Borell, Oxford University Press, 1999)."

BASE COUNT 348 a 105 c 176 g 257 t  
 ORIGIN

Query Match 4.4%; Score 167.4; DB 12; Length 886;  
 Best Local Similarity 53.8%; Pred. NO. 2.7e-39;  
 Matches 391; Conservative 0; Mismatches 331; Indels 5; Gaps 2;

QY 3056 tcgaacctcagagacattgaattcaggtatcgtcaacgtgattcggctcgaattctccgcg 3115  
 Db 110 TTGAATTTAAAGACATTTGTTTCAGATATCCCAACAGACCAGACAATTCGTCTGTGAAG 169  
 QY 3116 ggggttccatgacaatcgagccgaggaacaatttctagcttatgtggtcgtcttgcgtgtg 3175  
 Db 170 GTATTTCAATTTCAAAAGTAGAACAAGGAAAACCTCTTCCATTAGTAGAGCATCAGCATGTG 229  
 QY 3176 gcaagtcacgttgcgtcttgcggaaacgattctacgacccgcacctcgggcccgaattt 3235  
 Db 230 GTAAATCAACATCAGTTCAGTTGATTTGAAGATTTTATGATCCACACATCGAGATGTAT 289  
 QY 3236 catttgccacgagaaatattgcagaaatgtcgcgcgcgtctgtaccgcggccatattgtctt 3295  
 Db 290 TATTAGATGACATAATATCAAGATTTGAATATTTCAATTTCTTAAGAAAGTCAAAATTGGAA 349  
 QY 3296 tgggtcccaacgagccacacactttaccagaggtccggttcgcgagaaatgtga----cgttg 3351  
 Db 350 TGGTAGGACAAGAACCAAGTATTAATTTGCTGAAAGTCTTATGATTAATATTAGAGAGGAG 409  
 QY 3352 gccctcgaagccgaatattacgaagagctttgtcgaagagcgcctcccgcaagggccaatg 3411  
 Db 410 TACCTAAAGGAGTTGAAGTAACTAAATGAACAAATTTATGTCTGCTGCTAAAATGGCAATG 469  
 QY 3412 ctttggatttgcctctctttaccagaagcgtttgaaacgccttgcgcctcaacgaggg 3471  
 Db 470 CACATGACTTTATTTTCAGCAATGCCAGAGGATATACAC-AATGGTAGTGATAGAGT 528  
 QY 3472 atgcagtctccggggcgaacgacagcgaggtatgcgcctcgaagcattgattcgaaat 3531  
 Db 529 GCACAAATTTTCAGGAGGACAAAACAAAGAAATTCGCTATTGCCACGCTGATGATTAAGAAAT 588  
 QY 3532 ccaaaagcttgcctacttgacgaagcgacgtcagcctcgcacgcgaatcggaaagctctg 3591  
 Db 589 CCAAAAGTGTTATTAACGATGAAGCTACATCAGCAGCTTGATTCAGAAAAGTCAAAAGATT 648  
 QY 3592 gttaaacgtccctcgtatgaggaatccacagcgccgaacgaacgaatagcagtggtgcgaccca 3651  
 Db 649 GTACAAAGATGACCTTGACAAGCGCAAAAGGAAGAACACAACTGTAATTCACATAGA 708  
 QY 3652 ctttccactatcggaattgattgtttttgtttgttgcgaacggggaatccgcga 3711  
 Db 709 TTATCAACTATTCAAAATGCAGATCAAAATATGTTATTATGAGAGGAAGAAATTCAGAA 768  
 QY 3712 acgggcactcagcggaactcaacagcactgaaggaagatatattacgagatgtgttggca 3771  
 Db 769 AGAGGAACACATCAAGAGTTATTAGATTTCGAAGGATTTTATTATACACTTCTGTATGCA 828  
 QY 3772 caatctt 3778  
 Db 829 CAATTT 835

RESULT 4  
 AK014319  
 LOCUS Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length  
 DEFINITION enriched library, clone:32240JP09:ATP-binding cassette, sub-family  
 B (MDR/TAP), member 8, full insert sequence.  
 ACCESSION AK014319  
 VERSION AK014319.1 GI:12852089



Db 1714 TGGAGATATCCGATTTGGGAAGCTGGATGCTTCCGATGAAGAG----- 1757  
 QY 1466 tggagctgtgcaaaagcttgcaazagccagcaatggcgacgtgttcattaatgagcttc 1525  
 Db 1758 -----GTGTACACAGCTGCACGAGAGCAATGCCACGAGTTTCATCCACAGCTTCC 1809  
 QY 1526 cgaacgggtatgagactgaagtgtgcgagagcagccgagcccttgagtgagggtcaaacac 1585  
 Db 1810 CCGATGGCTAGACACTGTGTGGTGAGGGGGGCGACAACTTGTCTGGTGCCGAAGC 1869  
 QY 1586 aacgaattgcaatccagcaagtagtatcatcgatcccaagatccctgttactcgtatgaag 1645  
 Db 1870 AGCGCTAGCCATCGCAGCTGCCCTCATCAAGCAGCCGACAGCTGTGATGCTTGGACGAGG 1929  
 QY 1646 ctaccagcgcccttgacccgaagcgagaaagtgtgcagagagcccttgaaacagagtggt 1705  
 Db 1930 CCACCACTGCGCTAGATGCAAGATCCGAGAGGGTGTACAGAGGCGCCCTGGACCGGCCA 1989  
 QY 1706 ccaagagccgactactttagtctatgctcccaaaactgaccactatcaaaagtctggca 1765  
 Db 1990 GTGCTGGCGGACCGCTGTGGTCAATTCGCCACCGGCTTGTAGTACTGTCCGTGACGCCACT 2049  
 QY 1766 acatcgagcttccatccaggggaaaaatcgtcgagcaagcaagcaacacacgaattgatcg 1825  
 Db 2050 CCATCATTTGTCATGCGCAATGGCCAAAGTCTGTGAGGCTGGGACCCACGAAAGAACTCCTTA 2109  
 QY 1826 aattcggctgctcattagccgacactggtgcgtgcacagagccctcggggc 1874  
 Db 2110 AAAAGCGGGCTCTATTACAGACTTATCCGGAGACAAACCCCTGGACGC 2158

## RESULT 5

LOCUS A2532602 823 bp DNA linear GSS 03-NOV-2000  
 DEFINITION ENTCS0TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 genomic, DNA sequence.  
 ACCESSION A2532602  
 VERSION A2532602.1 GI:11088090  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica.  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 823)  
 AUTHORS Loftus,B., Van Aken,S. and Fraser,C.  
 TITLE Determination of clone end sequences from Entamoeba histolytica  
 HMI:IMSS sheared DNA library  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjl@loftusigr.org  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
 DNA library  
 Seq primer: M13-Reverse  
 Class: shotgun  
 High quality sequence start: 19  
 High quality sequence stop: 801.  
 Location/Qualifiers  
 1..823  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: pGEM1; Site: 1; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica  
 using a method described by Clark and Diamond (Clark,  
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
 method for isolate identification. Exp. Parasitol.  
 77:450.). The DNA was mechanically sheared to give a

## FEATURES

source

BASE COUNT 348 a 102 c 171 g 202 t  
 ORIGIN  
 tight size distribution (~2 Kb). The v + i method used for  
 the library construction is described in detail in Smith,  
 H.O. and Venter, J.C. (Making small insert libraries for  
 whole genome shotgun sequencing projects. In Genome  
 Sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Barrel, Oxford University Press, 1999)."

Query Match 4.3%; Score 164; DB 12; Length 823;  
 Best Local Similarity 54.2%; Pred. No. 2.8e-38;  
 Matches 385; Conservative 0; Mismatches 310; Indels 15; Gaps 2;  
 QY 1120 ggtcaaatgagatcccaaaacctggcatttgcctacccctcccgaccactctgcccgaagta 1179  
 Db 29 GGGAAATATTAGATTGGAGGATGTTCAATTTGTTTATCCAAAGACCATCTCATCATGTA 88  
 QY 1180 cttcagatttcaacttgacaattccagctggcaagacgagcgccctcgtcgtgcatca 1239  
 Db 89 TTAAAGGAATTTGACATTTGAATTAAGAAAGGAGAAACAATTGCANTTAGTTGGAGCATCA 148  
 QY 1240 ggtagcggcaaaagcacaatggtcggcttacttgaacggtgtgtctctgccagttcgggg 1299  
 Db 149 GGATGTGGGAGTCAACTACTATTTCATTAATCCAAAGAAATATATGAGCCAAATGGTGA 208  
 QY 1300 aggtattacttacttggtgtggaactggacaatacacaatgtgaaatggctgagagccgc 1359  
 Db 209 AGAGTAACATTTAGATGGAAGAGACATACGAGAAATGAATATCAATGGTTTAAAGAAATCA 268  
 QY 1360 attcgcctcgttcaacaggaacctgtgttctgttggtggcacaactcttccagaacattgcc 1419  
 Db 269 ATAGATTAGTTGGGACAGAACCCAGTGTGTTCAGAGAACAAATTCAGAAATATATTATG 328  
 QY 1420 aacggtttcattgatgcaacgagatctgcctcgcgaaacacaaatggagcttggcga 1479  
 Db 329 CTGGG-----AGCTAAAGAGGAAACACCAAGTAGAAGAGAGA-----TGATT 373  
 QY 1480 aaagcttgcgaagccagcaatggcgacgtgttcatttaagtgttcctcgaacggttatgag 1539  
 Db 374 GAATGTACTAATAATGGCAATGCACATGAATTTGTTTCTAACTTCAGAAAGATATGAC 433  
 QY 1540 actgaagtgtgcgagcggcggagccttgagtgaggtcaacgacaagaattgcaatc 1599  
 Db 434 ACAATAATTTGGAGAAAAGGAGCATTTATTCAGAGGACAAAACAAGAATTGCAATT 493  
 QY 1600 gcacgaagtatcatatcggtatcccaagatcctgttactcgtatgaagctaccagcgccct 1659  
 Db 494 GCACGTGCATTTGATTTCGAAACCCCATCTATTCTTCTTGTATGAAGCTACATCAGCACTT 553  
 QY 1660 gacccgaagcgaggaagaagtgttccagagggccttgaaccgagtggtcccaagaccgact 1719  
 Db 554 GATACACAAGTGAAGAGATTGTACAAAGAGCACTTCAAAAAGCATCTAAAGGAAGAAC 613  
 QY 1720 acttggctcattgcccacaactagccactgtcaaaagtctggaacatcgcagcatt 1779  
 Db 614 ACAATATTGTAGCACATAGACTAACAACTTTAGAAATGCAGATAAATATGTGTATT 673  
 QY 1780 tccaggggaaatcgtcagcaaggcacacaccacgaattgatcgaatt 1829  
 Db 674 CATCAAGGAGAAATTTATTGAACAAGGAAACATCAAGAATTAATGGATT 723

## RESULT 6

LOCUS AZ683753 947 bp DNA linear GSS 14-DEC-2000  
 DEFINITION ENT1196TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 genomic, DNA sequence.  
 ACCESSION AZ683753  
 VERSION AZ683753.1 GI:11820899  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota: Entamoebidae; Entamoeba.
1 (bases 1 to 947)
Loftus,B., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 40
High quality sequence stop: 778.
FEATURES
Location/Qualifiers
1..947
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).".
BASE COUNT
379 a 121 c 178 g 269 t
ORIGIN
Query Match 4.3%; Score 162.4; DB 12; Length 947;
Best Local Similarity 51.5%; Pred. No. 9.5e-38;
Matches 410; Conservative 0; Mismatches 371; Indels 15; Gaps 1;
QY 1036 ggtccgaattattcagtagtattcaaacaccacgcagctcgaccctctcgacccttct 1095
Db 1111111111111111111111111111111111111111111111111111111
81 GCTATTAAATATTACACAAATTGATAGATCCAGATATTGATGTCAGTCTATTGGA 140
QY 1096 ggaagcagccagagggctgcctaggtcacaattgagatcccaaaacctggcattgcctac 1155
Db 1111111111111111111111111111111111111111111111111111111
141 GGTGAATGTCCAACTGAGTGTATCGGAATATTAGATTGTGATGTTCAATTTGTTAT 200
QY 1156 cctcccgaccattcgcccaagctctcgagattccaaactgacaattccagctggcaag 1215
Db 1111111111111111111111111111111111111111111111111111111
201 CCAACAAGACCATCTCATCTATTAAAGGAATTCACCTTGAAATTAAGAAAGGAGAA 260
QY 1216 acacagccctcgtcgtgctacgtagcgcaaaagcaaatggcggcttaactgaa 1275
Db 1111111111111111111111111111111111111111111111111111111
261 ATAATTGCATTAGTTGGAGCATCAGATGTAGAAGCAATCTACTATTCAATTAATCCCA 320
QY 1276 cgtggtatctcccagtttcggggaggatattacttgatgggtgggaactgggacaatac 1335
Db 1111111111111111111111111111111111111111111111111111111
321 AGAAATATTCAACCAAAATGGTGGAAAGAGTAACATTAGATGGGAAAGAGAGATTG 380
QY 1336 aatgtgaattggctgagagccgcatctgcctctgctccacaggaacctggtgttcgt 1395
Db 1111111111111111111111111111111111111111111111111111111
381 AATATCAATGGTTAAGAAATCAATAGGAATAGTTGACCAAGAACCAAGTATTATTTC 440
QY 1396 ggcacaactctccagaacatgccacggtttcattgatgagcaacagagatgctctgc 1455
Db 1111111111111111111111111111111111111111111111111111111
441 GGAAACAAATTGGGAANAATATCATGCTTGACCTTAAGAAAGAGAA----- 485

```

RESULT 7  
 LOCUS AZ682350 871 bp DNA linear GSS 14-DEC-2000  
 DEFINITION ENTKB167P Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 genomic: DNA sequence.  
 ACCESSION AZ682350  
 VERSION AZ682350.1 GI:11819496  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica.  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 871)  
 Loftus,B., Van Aken,S. and Fraser,C.  
 Determination of clone end sequences from Entamoeba histolytica  
 HMI:IMSS sheared DNA library  
 Unpublished (2000)  
 CONTACT: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjloftus@tigr.org  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
 DNA library  
 Seq primer: M13-Forward  
 Class: shotgun  
 High quality sequence start: 16  
 High quality sequence stop: 860.  
 Location/Qualifiers  
 1..871  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba  
 histolytica Sheared DNA"  
 note="Vector: pHOSt1; Site\_1: Bst I; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica  
 using a method described by Clark and Diamond (Clark,  
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
 method for isolate identification. Exp. Parasitol.  
 77:450.). The DNA was mechanically sheared to give a  
 tight size distribution (~2 kb). The v + i method used for  
 the library construction is described in detail in Smith,  
 H.O. and Venter, J.C. (Making small insert libraries for  
 whole genome shotgun sequencing projects. In Genome  
 Sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Barrell, Oxford University Press, 1999).".











```

BH154857      899 bp      DNA      linear      GSS 24-SEP-2001
LOCUS      ENTRO47TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION      genomic, DNA sequence.
ACCSSION      BH154857
VERSION      BH154857.1 GI:15726574
KEYWORDS      Entamoeba histolytica.
SOURCE      Entamoeba histolytica.
ORGANISM      Entamoeba histolytica.
REFERENCE      1. (bases 1 to 899)
AUTHORS      Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
JOURNAL      HMI:IMSS sheared DNA library (2001)
COMMENT      Unpublished (2001)
              Contact: Brendan J Loftus
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0208
              Fax: 301 838 3543
              Email: bjlouftus@tigr.org
              Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
              DNA library
              Seq primer: M13-Forward
              Class: shotgun
              High quality sequence start: 42
              High quality sequence stop: 787.

FEATURES             source
     Location/Qualifiers
         1..899
             /organism="Entamoeba histolytica"
             /db_xref="taxon:5759"
             /clone_lib="Entamoeba histolytica Sheared DNA"
             /note="Vector: pROSI; Site.1: Bst I; Constructed at The
             Institute for Genomic Research (TIGR), Rockville, MD.
             Genomic DNA isolated from broth cultures of E. histolytica
             using a method described by Clark and Diamond (Clark,
             C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a
             method for isolate identification. Exp. Parasitol.
             77:450.). The DNA was mechanically sheared to give a
             tight size distribution (~2 kb). The v + i method used for
             the library construction is described in detail in Smith,
             H.O. and Venter, J.C. (Making small insert libraries for
             whole genome shotgun sequencing projects. In Genome
             Sequencing: A Practical Approach, eds. M. Vaudin and B.
             Barrell, Oxford University Press, 1999)."
```

```

BASE COUNT      358 a 119 c 170 g 252 t
ORIGIN

Query Match      4.0%; Score 150.4; DB 12; Length 899;
Best Local Similarity 51.0%; Pred. No. 4.4e-34;
Matches 418; Conservative 0; Mismatches 386; Indels 16; Gaps 2;

QY 1043 aattattcagtcattgacaaacccacgcagctgcacctctgcacctcttggaagc 1102
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 24 AAATATTACACTCTTTAGATCCAGATATTGATGTCATTCTATTGGAGTGGAAT 83
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1103 agccaggggctgctagtgcaattgagatccaaacccggcatttgcctaccctccc 1162
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84 GTCCAACGTGAGTGAATGGGAATATTAGATTTGATGATGTTCAATTTGTTATCCAA 143
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1163 gaccatctgccaaagtactcagatttcaacttgcaattccagctggcgaagcagcg 1222
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 144 GACCATCTCATGATTATAAAGGAATTGACCTTGAATTAAGAAAGGAGAAATAATTG 203
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1223 ccctcgtcggtgcatcaggtagcgcaaaagcacaatggtcggttacttgacgtggt 1282
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 204 CATTAGTGGAGCATCAGGATGAGAAAGTCAACTACTTCAATTAATCCAAAGAAAT 263
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1283 atctgccagtcggggagagatattacttgatgggttggaactgggacaatacaatgta 1342
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

DB 264 ATGAACCAAAATGGTGGAAAGAGTAACATTAGATGGGAAAGACATAAGAGAGACTTCAATATCA 323
QY 1343 aatgctgagaagccgcatctgctcttcaacaggaacctgtgttcttcggtgcacaa 1402
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 AATGGTTAAGAAATCAATAGGATAGTGGACAAAGAACAGTATATTTTCAGGAAACAA 383
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1403 tcttcagaaacattgcccaacggtttcatggtgagcgaacgagatctcctcgcgaaaaac 1462
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 384 TTGGAGAAATATCATGCTGTGAGCTAAAGAGGAGAA-----ACACTAA 428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1463 aaatgagcttgtcaaaaagcttgcgaagccagcaatggcgagctgttcatlaagagc 1522
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 429 GTAAGAAGAGAGATGATTGAATGTACTAAATGGCAATGACATGACTTTATTTCTAAAC 488
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1523 ttccgaacggttatgagactgaagtgtggcgagcgagcgaccttgagtgagggtcaac 1582
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 489 TTCCGAAGAGGATATGACACAAATAATTTGGAGAAAJAAGGAGCATTTATTCAGGAGGACAAA 548
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1583 gacaacgaattgcaatcgacacgaagtatcatatcgatcccaagatcctgttactcgtg 1642
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 549 AACAAAGAAATTCGAATTCGACGTGCATTGATTCGAAACCCCATCTATTCTTCTTGATG 608
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1643 agcttaccagcgcccttgaccccgaaagcgagagaaagtggtccaggaagcccttgaacgag 1702
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 609 AAGCTACATFCAGCACTTGTATACACAAAGTGAAAGATTGTACAAGAGAGCACTTGACAAAG 668
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1703 tgtccaaagacccgcacactactttggttcattgcccccaactagccactgtcaaaagtgtg 1762
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 669 CATCAAAAGGAAGAAACAACAAATTAFTGTAGGACATAGACTATCGACTATTCAAAATCGAG 728
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1763 gcaacatcgacgtcatcttcccccaggggaaatcgctcgagcagcacaacacacaaatga 1822
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 729 ATCAAAATGTTGTTATTATGAGAGGGAATAATAGTAGAACAAAGG-ACACATCAAGAAATAA 787
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1823 tgaattcggctgctcattacgcgcacgtggtgctgca 1862
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 TGGAGTTTGAAGGATTTTATTATACACTTGTCTATGCCACA 827
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
AZ687628/c      921 bp      DNA      linear      GSS 14-DEC-2000
LOCUS      ENTIR9STR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION      genomic, DNA sequence.
ACCSSION      AZ687628
VERSION      AZ687628.1 GI:11824774
KEYWORDS      GSS.
SOURCE      Entamoeba histolytica.
ORGANISM      Entamoeba histolytica.
REFERENCE      1. (bases 1 to 921)
AUTHORS      Loftus,B., Van Aken,S. and Fraser,C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
JOURNAL      HMI:IMSS sheared DNA library
COMMENT      Unpublished (2000)
              Contact: Brendan J Loftus
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0208
              Fax: 301 838 3543
              Email: bjlouftus@tigr.org
              Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
              DNA library
              Seq primer: M13-Reverse
              High quality sequence start: 25
              High quality sequence stop: 627.
              Location/Qualifiers
                  1..921
                      /organism="Entamoeba histolytica"
                      /strain="HMI:IMSS"
                      /db_xref="taxon:5759"

FEATURES             source
     Location/Qualifiers
```



```

QY 1446 tctgcctcgcgaacacacaaatggagcttgcacaaagcttgcaaaagccagcaatggcga 1505
Db 537 -----ACACTAAGTAAGAAAGAGAGATGATTGAATGCTACTAAATGGCAATGCACA 488
QY 1506 cgtgttatttaataatgagcttcgcgaacggttatgagactgaagttggcgaagcgcgcggagc 1565
Db 487 TGACTTTTATTCTTAACCTTCGGAAGGATATGACACAATAATTGGAGAAAAAGAGCATTT 428
QY 1566 cttgagtgagggtcaacgcacgaattgcaatcgacacgaagtatcatatcggatcccaa 1625
Db 427 ATTATCAGGAGGACAAAAACAAGATTTGCAATTCACCTGCGATTGATTCGAAACCCATC 368
QY 1626 gatcctgttactcgtgaagctacagcgcccttgacccgaagcggaagaaagtgtcca 1685
Db 367 TATTCTTCTTCTTGATGAAGCTACATCAGCACTTGATACACAAGTGAAAGATTGTACA 308
QY 1686 ggaggccttgaaccgagtggtccaaagccgcactacttgggtcattgcccacaaactagc 1745
Db 307 AGAAGCACTTGACAAAGCATCAAAAGGAAAGAAACAACAATTTATGTAGGACATAGACTATC 248
QY 1746 cactgtcaaaagtgcctggcaacatcgcaatcattcccgagggaacaaatcgtcgagcaagg 1805
Db 247 GACTATTCAAAATGCAGATCAAAATATGTGTATTATGAGAGGAAAAAATAGTAGACAAGG 188
QY 1806 cacacaccacgaattgatcgaattcggctgtcattacgcgcactgg 1852
Db 187 AACACATCAAGAAATTAATGGATTGAAAGGATTTATTATACACTTG 141

```

Search completed: April 26, 2002, 20:50:56  
 Job time: 5320 sec





[illegible][illegible]



Db 3142 tgtttttgactcagagcagaagtgttgacacatatgtatgtcagagttt--gcaggtta 3199  
Qy 2731 ctgtctaagatcatcacaagcttttggttggacgatgttttgggtttcactctatcacag 2790  
Db 3200 ccatacagaacactcttggagaaagcacacatcttggaaattacatttctcaccacag 3259  
Qy 2791 tcatgaggtttttgctattgcccctgggattttgtatgcagtcgataattggtcttcag 2850  
Db 3260 gcaatgatgtatttctccatgctggtgttccggttttgagcctacttgggtgacacat 3319  
Qy 2851 gtgagtcacgacaacactatttatcatctctcgtgggctttgtttgtccggtccaa 2910  
Db 3320 aaactcatgattgagatgtctctgttagttatttcagctgtgttcttgggtgccatg 3379  
Qy 2911 gcagcagccagctattgtcttactccagagttttaccagagctcgtgcgtgcgaac 2970  
Db 3380 gccgtggggaagctagcttctctgactatgcgaagccaaataatcagcagcc 3439  
Qy 2971 tatctctctggcgcgacattggaagccgacatccgcgaacgagaggaacaagaaa 3030  
Db 3440 cacatcatatgatctgaaaaaaccccttgattgacagctacagcacggaaggccta 3499  
Qy 3031 aaagcccaagtggtgagctgcctctgcagctcagagcattgaattcaggtatctcaaa 3090  
Db 3500 atccgaacacattggaagaaatgtcacatttggtaggtgttatcactatccacc 3559  
Qy 3091 cgtgattcggctcagttctccgcgggtttccatgacaatcgagccagggcaattgtta 3150  
Db 3560 cgacccgacatccagctctcagggagctgagcctggagggaagagccagacgctg 3619  
Qy 3151 gcttatgtggcgtctctgctgtggaagtcacagtcacagtcgtttgtcggagcagctc 3210  
Db 3620 gctctgtggcagcagtgctgtggaagagcacagtggttcagcgtctgagcggctc 3679  
Qy 3211 tagacccgacctcgggcgaatttctattgacacaggaatattgcgagaatgtcgccg 3270  
Db 3680 taagacccttgcagggaaaagtgcctgtgtggtggaagaaataaagcagctgaattgt 3739  
Qy 3271 cgtttgacagcggccatagtcttctgtccaaagaggaaccacacattaccagggctcc 3330  
Db 3740 cagtgtcctcgagcacctcgggcatcgtgtccagggagccatcctgttgcagtcagc 3799  
Qy 3331 gtctgcgagaa-tgtgacgttgccctcgaaagccgaatattacgaagagctttgtcaag 3389  
Db 3800 atgtctgagaacattgctcatgagacacagccgggtggtgtcacaggaagagattgtg 3859  
Qy 3390 acgctctccgcagagcgaatgtttgattgttctcatctcttaccagaaggtttgaa 3449  
Db 3860 agggcagaagagggccacatacatcgtctcatcgtgagctactgctcaataaataagc 3919  
Qy 3450 acgcttgcggtcacaagggatgcagttctccgcggggaacacgacagcggatgccat 3509  
Db 3920 actaaatgagagaca-aaggaactcagctctctgtggtgccaagaacacgattgccat 3978  
Qy 3510 cgaaagagcattgattcgaaatcacaagctgtgtctactgtacgagcagcagcagccct 3569  
Db 3979 agctgtgctctgttagacagcctcatatttgttttggatgaagccacagtcagctct 4038  
Qy 3570 cgaaacgaatcgaagctctgttgaagctgccttcagatgagggatccacagccgaac 3629  
Db 4039 ggatacagaagtgaaaggtttgtccaagaagccctgtgacaagccagagagggccgac 4098  
Qy 3630 gacaatagcagtggtgcacagcacttccactattcggaaatgtgtgtattttgtatt 3689  
Db 4099 ctgcatgtgtatgtctaccgctgtccacatccagaatcgagacttaattagtggttt 4158  
Qy 3690 tgcacagggagaaatcgcgaacagggcactcactcagcggaaactacacagcagtaggag 3749  
Db 4159 tcagaaatggcagatgcaagagcatggcagcagcagcagcagcagcagcagcagcag 4218  
Qy 3750 atatt 3754  
Db 4219 ctatt 4223

RESULT 2  
US-10-097-340-1  
; Sequence 1, Application US/10097340  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; PRIOR FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 4643  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-097-340-1

Query Match 6.8%; Score 259; DB 6; Length 4643;  
Best Local Similarity 44.9%; Pred. No. 1.8e-70;  
Matches 1537; Conservative 1; Mismatches 1821; Indels 66; Gaps 12;

Qy 360 catttcagctatcacgaacccactcgtcgcagcagtggttcttgaatgcacctgcg 419  
Db 835 cctggcagctggagagacaaatacacaaaattagaaaaacagtttttctaatgataaagcg 894  
Qy 420 gcaagaggtctggcatttcgcacagcaggaatggagcaatccactcgcactcactac 479  
Db 895 acaggagatagctggttggatgtgcagcagtggtgggagcttaacacccgactacaga 954  
Qy 480 caatggcaaccgtatatacaaacaggtattgcgcagaaattggtctttaccgtgcaggcact 539  
Db 955 tgaatgctccaagattaatgaaggaattggtagacaaaattggaattgtcttcttcagtaac 1014  
Qy 540 ttcattgtttcttctgattgtgtgtcgtttgtgtcgttctcagtggaagctagctttaa 599  
Db 1015 ggcaaatcttcaactgggttttatagtaggatttacacgtgtgttggaagctaacacctgt 1074  
Qy 600 caccatgctcgtccctgcctcctatttctgtgtcaccggcactcgcagacaaatgatgc 659  
Db 1075 gatttggcactcagctcctgtctgttgactgtcagctgtctgtgggcaagataactatc 1134



```

Db 3260 gcaatgctgtattttctctatgctggatgtttcccggttttggagcctacttgggtggccat 3319
Qy 2851 gtgagtcagacacaactcaattttatatcatcttcggtggcggttttgggtccggtccaa 2910
Db 3320 aaactcatgagcttggagatgtctgttagtatatttcagctgttcttcttggcgcatg 3379
Qy 2911 gcagagccacagatatttggcttactccacagtgattttaccagagctcggtcggtcgaaac 2970
Db 3380 gcegtgggcaagtcagtcagtcatttctcctgactgctgacatggccaaagccaaatcatcagcgc 3439
Qy 2971 tatactctcggctcgagcagcattgaagcgcagccaccatccgcgaacgcgggaggaacaagaaa 3030
Db 3440 cacaatcatgatcattgaaaaaaccccttggattgacagctacagcagcaggaaggccta 3499
Qy 3031 aaagggccagctgggtgatgccctctgcgacctcgagacatgttgagacattgaattcaggtatcgtcaa 3090
Db 3500 atgcgcaacacattgggaagaaatgtcacatttggtagatttggtagatttattcaactatcccaac 3559
Qy 3091 cgtgattcggctcagtgctcccggtgttccatgacaatcgagccagagcacaatttcta 3150
Db 3560 cgaccgacatccacgtgcttcaggagactgagcctggaggtgaagaaggccagacgctg 3619
Qy 3151 gcttatgtggcgcttctgctgtggaagtcacagcttgatcgcttctgtcggaacgattc 3210
Db 3620 gctctgtgtggcagcagctggctgtgggaagcagcagtggtccagctcctggagcggctc 3679
Qy 3211 taqcaaccgacctcgcgccgaatttcttgacacagagaaatattgcagaaatgtcgcg 3270
Db 3680 taqcaacccttggcagggaaagtgtctgtgtgagcgaagaataaaagcagactgaatgtt 3739
Qy 3271 cgcttgcagcgccatagtcttcttgggtccaaacaggaacccacactttaccagaggctcc 3330
Db 3740 cagtggtccgagacacacctggcgatcgtgtccagagagccatcctgtttagtcgagc 3799
Qy 3331 gttcgagagaa-tgtgactgtggccctcgaaagccgaattatcagaagagctttgtcaagg 3389
Db 3800 attgctgagaacattgctatggagacacagccgggtgtgtcacaaggaagagattgtg 3859
Qy 3390 acgcttcccgcaaggccaatgtcttggatttggatctctcttaccagagggttttga 3449
Db 3860 agggcagcaaaaggggccacacatacatgccttccatcgagtcactgactgactgactgactg 3919
Qy 3450 acgcttgcggtcaacgagggatgcaggttctccggtccggtcggaacgagcggatcgccat 3509
Db 3920 actaaagttagagaca-aaggaaactcagctctctgtgtggccagaaacacacgcatggcat 3978
Qy 3510 cgaagagcatgtattcgaataatcgaagctgttgcacttgcaggaagcagcgtcagccct 3569
Db 3979 agctcgtgacctttagacagagcctcatatttcttggatgaagccacgtcagctct 4038
Qy 3570 cgacacgcaatcggaaagctcgtgttcaagctgcctcgatgaggtcaccagagccgaac 3629
Db 4039 ggatacagaagtgaaaagggtgtccaaagaagccctggacaaagccagagaagccgcac 4098
Qy 3630 gacaatagcagtggtgcagccagacttccactatttcggaatgttgatatttttggtt 3689
Db 4099 ctgcatgtgtattgtctcaccgctgtccaccatccagaatgcagacttaagtggtgtt 4158
Qy 3690 tgcacaagggagaaatccggaacggaactcagcgggaactacaagactgagagaag 3749
Db 4159 tcagaaatggcagatgaaggagatggcagcagcatcagcagctgctggtgcacagaaaggcat 4218
Qy 3750 atatt 3754
Db 4219 ctatt 4223

```

RESULT 3

US-09-540-209B-1731  
; Sequence 1731, Application US/09540209B  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1731
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-1731

```

Query Match 3-5%: Score 133.2; DB 5; Length 1839;  
Best Local Similarity 51.0%; Pred No 2.2e-31;  
Matches 368; Conservative 0; Mismatches 348; Indels 6; Gaps 2;

```

Qy 3011 aaacggaggagaaacaaagaaagccagtggtggtgagtcctgtgcacctgcagctgcagacaa 3070
Db 1073 agatgtggagattccgaaccgaaacccgttgatgggttggaaagagcaagtcaggtcca 1132
Qy 3071 ttgaattcaggtatcgtcaacgtagctcgtcaggttcccggttcccgaggttccatgacaa 3130
Db 1133 agatatttttccagttatgaggaagaagaagcacttcaacatattaaacctgactg 1192
Qy 3131 tcagccagagacaattttagcttatgtggcgcttctgtgctgtggcgaagtcaactgtga 3190
Db 1193 taaccaaggtaaaacacctgttgggtggcagtcggttcaggaagactacgttgg 1252
Qy 3191 tcgcttggtcggaacgattatagacccgacctcgccgaatttccggaatttgcacacagaa 3250
Db 1253 tagattcatgtccacgctatcacgattgacgaaggaagcagattcactatagacggaatca 1312
Qy 3251 atattgcagaaatgtgcccgctgtaccgcggcccatatgtcttcttgggtccaaacaggaac 3310
Db 1313 atacaagagattgtctgtatctgtattgtgcagcctgataggaaaagtcaatcaggag 1372
Qy 3311 ccacactttccaaagctcgttcgcgaagtgtgacgttggcctcgcagcgcggaatatt 3370
Db 1373 ctatttgcgaatgatacttctcaataatattgttctggcggttgagaacgcgacta 1432
Qy 3371 cagaagagcttgcgaaggagccttcccgcaaggcccaatgcttggatttgcactctc 3430
Db 1433 tggagcaggtgatgta----ggccgcaaaatagcacaatgcacacgattttatcattgga 1487
Qy 3431 ttaccacagagctttaaagccttcggtcgaacgaggtatgcaatctcccgcgcggc 3490
Db 1488 gaaagaagcgggtaccataactaaatctgg-tgaccggggcagcaagctctccgggggccc 1546
Qy 3491 aacgacagcggtatcgctacgcaagagcattgattcgaataatccaaagctgttgcactctg 3550
Db 1547 aacgcccagcgcacgacagcagccgtgccaatcttgaagaatccctccatcttgcattctg 1606
Qy 3551 acgaagcagctcagcctcgacacgaatcggaacgtctggttcaagctgcctcctcatg 3610
Db 1607 acgaggtacttcgcttggatgactgagtcagagcgttgggtacaggaagcgtcgagc 1666
Qy 3611 aggcacacagcgcgaacgaacatagcagtgccgcaccgacatttccactatttcggaatg 3670
Db 1667 gtttgatgaacacgctacgacgattgtctatttgcacccgctctctaccataaaaaatg 1726
Qy 3671 ttgattgtatttttggatttgcgaacggggagaaatcgccggaacgggagcactacgcggaaac 3730
Db 1727 cagatgaatttgtgtctgtacgaaggggaaattgtagaacgcgggaaacacacgaagaat 1786
Qy 3731 ta 3732
Db 1787 ta 1788

```

RESULT 4

US-09-540-209B-3205  
; Sequence 3205, Application US/09540209B

```
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3205
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-3205
```

```
Query Match      2.6%; Score 100.2; DB 5; Length 2223;
Best Local Similarity 49.6%; Pred. No. 5.1e-21;
Matches 326; Conservative 0; Mismatches 313; Indels 18; Gaps 2;

QY 1173 ccaagtaactcagatttcaacttgacaattccagctcgtggaagacagcagcgcctcgtcgg 1232
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1536 caagattagacaataataattagatataccgagggaaagattacagctatcgtagg 1595

QY 1233 tgcatacagtagcgcgaagacacaaatggtcgttacttgaaacgttggttatcgtccag 1292
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1596 ggaagcggagcgttaaaacacattactaaaattgttactcaggtttataaacaac 1655

QY 1293 ttcggggaggattactctgatgggttggaactggggaacaatacatatgtgaaatgctcag 1352
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1656 aaatggtgaaattgaagtcggggaggtccattagataatcgtatttatcgttggcg 1715

QY 1353 agcgcgcatccctcgttcaacaggaacctgtgttctgtgacaaatcttccagaa 1412
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1716 taatggtgcgtgcgtattacaagatggttaaaacttttaagacactatctataaa 1775

QY 1413 cattgccaacggttctatgtagcaacgagatcgtcctgcgcgaacaaataatggagct 1472
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1776 tatcaca-----ttggaagatgaagaatgaacgtataatcaaaacaaactagtgaagc 1829

QY 1473 tctgcaaaaagcttgcaagccagcaatggcgacgtgttcattaatgagctccgaaacgg 1532
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1830 gattcaactggca-----aacgcagaaaattcattaatatgctgccttttaa 1877

QY 1533 ttatgagactgaagtggcagcgcggcgttgagtcgtgagtgcaacgacaacgaat 1592
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1878 gctatacctccacttgaactaatggttcaggatgaagtcgaaggcgaacgaacgtat 1937

QY 1593 tsgaatcgcagcaagtatacatcgtatcccaagatcctgttactcgtatgaagctaccag 1652
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1938 actaatagtagggcgatatacaaaaacccagactctatcttctggatgaagcgacaaa 1997

QY 1653 cgcctctgaccgaaaggcggagaaagtgtccagagggccttgaaccgagtgctccaaaga 1712
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1998 ctctctgacacgaataatgaaaaacagatttcccaaaatttggaaacattttagaagg 2057

QY 1713 ccgcactacttggctcattgccacaaactagcactctcaaaagtctggtggaacatcgc 1772
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 2058 aaaaaccgcattgtaatagcgcacccagattgaagctatgataaaatgcacataatagt 2117

QY 1773 agtcatttccaggggaaaatcgtcagcaaggcgcacacccagaaattgacgaatt 1829
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 2118 ggtcatgaaaaggaaaattgtggaacagggactcatcaagaactataaattt 2174
```

```
RESULT
US-09-540-209B-1067
; Sequence 1067, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
```

```
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1067
; LENGTH: 648
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-1067
```

```
Query Match      2.6%; Score 98.8; DB 5; Length 648;
Best Local Similarity 55.6%; Pred. No. 7e-21;
Matches 190; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
```

```
QY 1482 agcttgcaaacgacgaatgagcagctgttcaatgaagcttcggaacggttatgagac 1541
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 240 agcggccaaaagcctgtgttcacgactttatcatgcatctgcacatggttaagatcac 299

QY 1542 tgaacttgccagcagcgcgagccttgagtggaagctcaacgacaacgaattgcaatcgc 1601
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 300 aatgctggggagaggggctgtacgttatcaggtgtgtgaaaaaacacggctttccattgc 359

QY 1602 acgaagtatcatatcggtatcccaagatcctgttactcgtatgaagctaccagcgccttga 1661
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 360 gctgtccatgctgaaagcgcacagatcgttctgtgacgagggcaactgcttcgttga 419

QY 1662 ccggaaggcggagaagaagtggtccagagccttgaaacggagtggtccaaagacgcactac 1721
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 420 tcccgagaacgaagtagagatcacagaaggtatcgtatcgttgattaaaggcgaacggt 479

QY 1722 ttgtgtatggcccaacaaacagccactgtcaaaagtgctggaacacatgcgcagtcatttc 1781
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 480 cattgttatcgccactcgtctcaagacaataatggtggcgcacacactgtgtttatttc 539

QY 1782 ccaggggaaaaatcgtcggagcagggcacacacacgaattgat 1823
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 540 cgtatggaaaagtggaagacaaggtcacgcatttcgggaattgat 581
```

```
RESULT
US-09-540-209B-966
; Sequence 966, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 966
; LENGTH: 1980
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-966
```

```
Query Match      2.6%; Score 98.2; DB 5; Length 1980;
Best Local Similarity 55.8%; Pred. No. 2e-20;
Matches 187; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
```

```
QY 1495 agcaatggcgacgtgttcaatgaagcttcggaacggttatgagactgaagttgacgag 1554
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1603 accgtgcccattcgttcatcaaacgcttgcgcaaggttagcatcttctctcgaaaac 1662

QY 1555 cgagccgagccttgagtgagggtcaacgacaagaatgcaatgcgcgaagtatcata 1614
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1663 gacggagccaacctgagccaagggcagcgtattgaaacttgcctgcccagtg 1722

QY 1615 tcgatatcccaagatcctgttactgatgaagctaccagcgccttgacccgaagcggag 1674
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1723 gccgatccgcccactcctgttctggaacgacgaacgaacacatagatcacgcagtgaa 1782

QY 1675 aaagtgtccaggaggccttgaaccgagtggtccaaagaccgcactacttggctattgcc 1734
```

Db 1783 atctgaccagcggggttgacactgttgatgaagagcaccagcctgatcatgcc 1842  
Qy 1735 cacaactagcactgtcaaaagtctgggaacatcgacgtatctccaggggaaatc 1794  
Db 1843 caccgctgtcaagtcgcaatgcagataccatctgtgtagcagcaggaaatc 1902  
Qy 1795 gtcgacgaagcacacaccaggaattgatgaatt 1829  
Db 1903 atcgaagcagtcacatcaggaaattacttgatt 1937

## RESULT 7

US-10-102-806-308  
; Sequence 308, Application US/10102806

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA103PlC1

; CURRENT APPLICATION NUMBER: US/10/102,806

; PRIOR FILING DATE: 2002-03-22

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/03881

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 846

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 308

; LENGTH: 2171

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1248)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc.feature

; LOCATION: (2162)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc.feature

; LOCATION: (2166)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc.feature

; LOCATION: (2168)

; OTHER INFORMATION: n equals a,t,g, or c

US-10-102-806-308

Query Match 2.5%; Score 93.4; DB 6; Length 2171;

Best Local Similarity 49.6%; Pred No. 6.7e-19;

Matches 287; Conservative 0; Mismatches 271; Indels 21; Gaps 1;

Qy 1120 ggtcaaatgtagtccaaaacctggcatttgcctaccctcccgaccatctgccaagta 1179

Db 1585 ggccctgtccagttccaaagatgtctcttgcctaccacccgcagatgtcttagtg 1644

Qy 1180 ctctcagatttcaacttgacaattccagctggcaagacgagcgcctctgctgtagatca 1239

Db 1645 ctacaggggctgacattcacctacgacctggcgaggtgacgagcgtggtgaccat 1704

Qy 1240 ggtagcggcaaaagacaaatggtcgcttacttgaaacggtggtatctgcccagttcggg 1299

Db 1705 gggctgggaagagcacagtgctgctcgtcgagaaatctgaccagcccacgggga 1764

Qy 1300 aggatattactgtggttgaactgggaataataatgataatgaaatggctgagaagccgc 1359

Db 1765 cagctgctgttgtaggggaagccctctcccaatatgagcaccgctacctgcacaggcag 1824

Qy 1360 attcgccttgtaacaggaacctgtgttttctgtggcaaatcttccagaaactgcc 1419

Db 1825 gtcgctgcagtggggcaagagccacaggtatttgggaagagtgcttccaaataattgcc 1884

Qy 1420 aacggttttcattgtagtagcaacagagatctgcctcggaaaaaataatggagcttgtgcaa 1479

Db 1885 tatggctgaccagagaagccaact-----atggagggaataacaca 1923

Qy 1480 aaagctgtcaaaagccagcaaatgacgactgtttcattaatgagcttccgaacggttatgag 1539

Db 1924 gctgcgcagtaaaagtctggggcccatagtttcattctctggacctccctcagggtatgac 1983

Qy 1540 actgaagttggcagcagcagcagcagccttgagtgagatcaacacacgaattgcaatc 1599

Db 1984 acagaggtagacgaggtggtggagccagctgtcaggggttcagcagcagcaggtggcgttg 2043

Qy 1600 gcacgaagtattcatatcggtatcccaagaatcctcttactcgatgaagctaccagcgccctt 1659

Db 2044 gcccgagcattgataccgaaacccgtgttacttctcctggatgatgccaccagtgccctg 2103

Qy 1660 gacccgaagcggagaagtggtccagcagcagcgttgaac 1698

Db 2104 gatgcaaacagccagttacaggtggagcagctccctgtac 2142

## RESULT 8

US-09-540-209B-527

; Sequence 527, Application US/09540209B

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES

; FILE REFERENCE: 2709,1001-001

; CURRENT APPLICATION NUMBER: US/09/540,209B

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 10444

; SEQ ID NO 527

; LENGTH: 2208

; TYPE: DNA

; ORGANISM: B.fragilis

US-09-540-209B-527

Query Match 2.4%; Score 92; DB 5; Length 2208;

Best Local Similarity 48.4%; Pred No. 1.9e-18;

Matches 337; Conservative 0; Mismatches 335; Indels 24; Gaps 2;

Qy 1168 totgcccgaagtactctcagagatttcaacttgacaattccagctggcagacacagcggccctc 1227

Db 1525 tctccttggctgcgtataatttacttacaataaccggaaaaataaataacacggcaatt 1584

Qy 1228 gtcggtgcatcaggtagcggcaaaagacaaatggtcgcttacttgaaacggtggtatctg 1287

Db 1585 gtgggaggaagtgtagtgtaagtaactcttctgaaactatggttcggtttataaag 1644

Qy 1288 ccagttcggggaggatataattacttgatgggttgggaactgggacaatacaatgtgaaatgg 1347

Db 1645 ccagccatgggaaataaaatggacaagatgaatgaagtgcattaaactacgcccac 1704

Qy 1348 ctgagaagcgcgcttcctcgtcttcaacagggaacctgtgtgttctcgtggcacaattctc 1407

Db 1705 tggagaacaatgtgtgggtgtgtaatgcaagatggaaaaatattcagtgataccattcg 1764

Qy 1408 cagaacaattgccaacggtttcattgatacagcagagatctgcctcgcgaaaaaacaatg 1467

Db 1765 aataaatgtg-----taitagatgatgaac-----aataataattat 1800

Qy 1468 gagcttgtcaaaaagcttgcaaacccagcaatggcgagctgttcaattaatgagcttcog 1527

Db 1801 acgctgtgggaagatttctcgtatcgtcagattgaggatgagataaacgcgatgcct 1860

Qy 1528 aacggttatgagatgaagtgtgcgagcagcggagccttgagtgaggtgaggtcaacagcaa 1587

Db 1861 aaggggtttgaaaacaccattggagaacccgacgcgggttgagtgagagacaaagcag 1920

Qy 1588 cgaatgtcaatgcagcagagtatcatatcggatcccaagctctgttactcgatgaagct 1647

Db	1921	cgtttggtgattgctcgtgcgctgtatcgggataccgaaattctcttttagacgaagcc	1980
Qy	1648	accagcgccttgaccgcgaagcgagagaaagtgcacgagggccttgaaacgagtgcc	1707
Db	1981	acaaactcttggtatcaataatgaacgaaaaattgtgaatgccttgaacaatgcatt	2040
Qy	1708	aaagaccgcgactatttggctcatttcccccacaaactaagccacttcaaaagtctggcaac	1767
Db	2041	gaacagcgactctgttcttatttctcaccagggcttagtaccattcgttaatgctgacaa	2100
Qy	1768	atcgcgagtcatttccacgggggaaaaatcgtcagcaaggcacacacagaaatgatcgaa	1827
Db	2101	artgtggtgtgcgaaaaaggttttatctgttagacgcgaactaatgaaattatgatgag	2160
Qy	1828	ttcgcgtgtcattacgcgcgactggtgcgtgcacag	1863
Db	2161	aaaaagggggactatttatttgattggtgtctcttcacag	2196

## RESULT 9

```

US-10-105-299-238
/ Sequence 238, Application US/10105299
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, et. al
/ TITLE OF INVENTION: Human Secreted Proteins
/ FILE REFERENCE: PS950
/ CURRENT APPLICATION NUMBER: US/10/105,299
/ CURRENT FILING DATE: 2002-03-26
/ NUMBER OF SEQ ID NOS: 15197
/ Prior Application removed - See File Wrapper
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 238
/ LENGTH: 721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-105-299-238

```

Query Match 2.2%; Score 82; DB 6; Length 721;  
Best Local Similarity 52.0%; Pred. No. 1.3e-15;  
Matches 184; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY	1512	caftaatgagctctccgaacggttatgagactgaagtgtgcgagcagcgagcgttgag	1571
Pb	67	caataatggcttctccctgaaggttacaggaacacagctggcgagcgggactgaagctgag	126
QY	1572	ttagaggtcaacgcacgaacgaattgcaatcgacgaagtatcatatcgattcccaagatacct	1631
Pb	127	cggcggggagaagcagcgcgctcgccattgcccgcaccatctcaaggtctcgggcatcat	186
QY	1632	gttactcgatgaagcttaccagcgcccttgaccccgaaagcggaagaagtgttccaggaaggc	1691
Pb	187	tctgctggaagagcgaactcgcgctggatcacatcaatgagagggccatccaggtctc	246
QY	1692	cttgaaacgcgagtggtccaaagacgcgaactactttggtcattgcccacaaactgcccactgt	1751
Pb	247	tctggccaaagtctgycgaacgcgaccacacatcgtagtggcacacaggctctcaactgt	306
QY	1752	caaaagtcgcggcaacatcgactgcatcttccaggggaaaactgtcgagcaaggcacaca	1811
Pb	307	ggtaatgctgacacagatctctgcatcaagaatggtcgtcgctggagaggggacgaca	366
QY	1812	ccacgaattgatcgaattggctgtcattagccgcgaactggtgcgtgcacaggaa	1865
Pb	367	cgaggtctgtgtcccgaggtgggggtatagtgcacattatgcaactatgcaactctcaaca	420

## RESULT

US-09-540-210B-34545  
; Sequence 34545, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Deleceane, Angelo M.

```
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 34545
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: rat00136905
US-09-540-210B-34545
```

```
Query Match 2.1%; Score 79; DB 5; Length 304;
Best Local Similarity 61.4%; Pred. No. 7.1e-15;
Matches 127; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1617 ggatccaagatctgttactcgtatgaagctaccagcgccttgaccgaaagcgagaaa 1676
Db 1 gaaccccaaaattctctcgtatgaagcaaccgacgctggtgacgaaacgagca 60
QY 1677 agtggtccagagccttgaccgagtggtccaaagaccgcactactttgttcattgccca 1736
Db 61 cctggtccagagcgcttgaccggtgatggagggaagaaacgctgtgatcatagccca 120
QY 1737 caaatgacacttcaaaagtctggcaacatcgacgactcatttcccaagggaataatcgt 1796
Db 121 cgcctctccaccattgaagcgcgaattctcgtgcgctccttaaccagaagaaatcgt 180
QY 1797 cgagcaaggcaccacacgaattgat 1823
Db 181 tgaacatgggacacatgaggagctgct 207
```

```
RESULT 11
US-10-106-698-2063
; Sequence 2063, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
```

```
; SEQ ID NO 2063
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (57)..(57)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2063

Query Match 2.0%; Score 75.4; DB 6; Length 807;
Best Local Similarity 56.1%; Pred. No. 1.6e-13;
Matches 142; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 1573 ggaggtccaacgacacgaatttcgaacgaagatcatatcggatcccaagaatcgtg 1632
Db 259 ggagggagaaaagcaaaagatagcaatttgaagaccatttgaagccccccagtcata 318
QY 1633 ttactcgtatgaagctaccagcgccttgaccgaaagcgagaaagtgtgtccagagacc 1692
Db 319 ctctatgatgaagctacttcctcgttagattcgattgaagagactattcttggcgcc 378
QY 1693 ttgaaccgagtgctccaaagaccgcactactttgttcattgcccaaaactagccactgtc 1752
Db 379 atgaaggatgtgtgtccaaacacagaaactttatttcttcattgcacacagattgtcaacagt 438
QY 1753 aaaagtgtggcaacatcgactcatttccaggggaaaatcgtcgagcaagcacacac 1812
Db 439 gttgatgcagatgaataatcatttctgttgatcagggtagccgaactggtgaccacac 498
QY 1813 cacyaattgatcgt 1825
Db 499 catggtttgtctg 511
```

```
RESULT 12
US-09-540-209B-3403
; Sequence 3403, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3403
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: B.fragilis
US-09-540-209B-3403
```

```
Query Match 1.9%; Score 72.2; DB 5; Length 2223;
Best Local Similarity 45.9%; Pred. No. 2.9e-12;
Matches 322; Conservative 0; Mismatches 373; Indels 6; Gaps 2;

QY 3062 tcagagacattgaattcaggtatcgttcaacgtgattcggctcgagcttccgcggggttt 3121
Db 1499 tcagaaatctcgttccagtattgacggtgacgctcgtgactatgtgtgaatgatgta 1558
QY 3122 ccattgacaatcgagccaggaacaattttagcttatgtggcgcttctggtgtggtgcaagt 3181
Db 1559 acctgaacatacccggaacaaaggtgacggtgattgtgggtgagcggtagcggaaga 1618
QY 3182 caacgttgatcgtttgtcggaacgattctacgacacactcgccgacatttcattttg 3241
Db 1619 ctacgctgatcaagctgctgtgcttttactcccaataaggggggataatcaagaatg 1678
QY 3242 cacacgagaatattgcagaaatgtcgcgcgcttgtaaccgcgcacatatgtcttggatcc 3301
Db 1679 gtgagacccctctggtatgtggtataatccgcacactctggtgcgaagagcggttcgtgta 1738
```

Db	1312	caagaagaactcgaagtggcagccctgc-----gccgaagccatgagttcatctgct	1366
Qy	3430	ctttaccagaagcgtttgaaacgcccttcggtccaacagggatcgagttccgcgcggg	3489
Db	1367	cgttgectgcggatacgcac-acfgctggacgggggtggaaatcaacctttcgcgcgga	1425
Qy	3490	caacgacagcggatcgccatcgcaagagcatttgattcgaaatccaaaagctgtctactt	3549
Db	1426	cagaagcaacgcgatactgcacgcaggccttcgcgcgtataaccacgatatattatactg	1485
Qy	3550	gacgaagcagctcagccctcgacacgcgaatcggaaacgtctggttcagagctgccctcgat	3609
Db	1486	gacgacgataccagtcgctagagactctgaacgcggaactcgtatccggaacaactgaaat	1545
Qy	3610	gaggcatccacgacgcgaacgacaaatgacgtggcgccacgcctctccactattcggaaat	3669
Db	1546	gcctgctgcgggataccacgcgggtctatcatcaccacgcatacacacatgcgaatcg	1605
Qy	3670	gttgatgtattttgtgttgcgaacgggagaatcgccgaacgggcactcaacgcggaa	3729
Db	1606	gccaaacgggtcattctctctgacgacgggggagatgaatccatctcggcaccccgaaagag	1665
Qy	3730	ctacaacgactgagagaaagatatcagagatgtttggcaca	3773
Db	1666	ttgttaaacacgtccgaataatgtatccgaaataataattactcaca	1709

```

US-09-975-254-14965
? Sequence 14965, Application US/09975254
? GENERAL INFORMATION:
? APPLICANT: BYrum, Joseph R.
? APPLICANT: Heck, Gregory R.
? APPLICANT: La Rosa, Thomas J.
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: plants
? FILE REFERENCE: 38-21(153091B)
? CURRENT APPLICATION NUMBER: US/09/975,254
? CURRENT FILING DATE: 2001-10-12
? PRIOR APPLICATION NUMBER: US/09/263,191
? PRIOR FILING DATE: 1999-03-05
? NUMBER OF SEQ ID NOS: 31255
? SEQ ID NO 14965
? LENGTH: 237
? TYPE: DNA
? ORGANISM: Glycine max
? OTHER INFORMATION: Clone ID: 700909087H1
US-09-975-254-14965

```

Query Match	1.8%	Score 68.6	DB 5	Length 237
Best Local Similarity	59.4%	Pred. No. 1.1e-11		
Matches 133	Conservative	0	Mismatches	Indels 1
Gaps	1			
Qy 1181	ttcggagattcaactgacattccagctgtgccaagacgagccctcgtcgttgccatcag	1240		
Db	2 ttcgagacatcagctgactatccattctgcgacagcagtggtctgttgggtgaagctg	61		
Qy 1241	gtagcggcaaaagc-acaaatgctcggtctacttgaacggttggtatctgccaggtctcggg	1299		
Db	62 gaagtggaaaaacacacagctattgcctgtgtgcaagattttatactcgtatcaggc	121		
Qy 1300	aggatattacttgatgggttggaaactggggacaaacaaatgtgaaaaggctggaagccgc	1359		
Db	122 caaatcacacttgatgggatagaatactgtgagttacagctcaangtgggtaangcngcag	181		
Qy 1360	attcgcctcgttccacaggaccctgtgtgttctggtggcaaat	1403		
Db	182 ataggccttgtaagccaagaccagttttgttcaatdaaactat	225		





\_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 01:34:27 ; Search time 8172.36 seconds  
(without alignments)  
10038.659 Million cell updates/sec

Title: US-09-882-694A-10

Perfect score: 3792

Sequence: 1 atgscagatgaatcgagaa.....aatcttagacaagcatga 3792

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_NA\_Main:\*

- 1: /cgn2\_6/ptodata/2/pna/pna/PCTUS\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*
- 8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*
- 23: /cgn2\_6/ptodata/2/pna/US095E\_COMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*
- 25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*
- 26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*
- 27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*
- 28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*
- 29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*
- 30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*
- 31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*
- 32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*
- 33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*
- 34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq.\*
- 35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq.\*
- 36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq.\*
- 37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq.\*
- 38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq.\*
- 39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*
- 40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*
- 41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*
- 42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*
- 43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*

- 44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*
- 45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*
- 46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*
- 47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*
- 48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*
- 49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*
- 50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*
- 51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*
- 52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*
- 53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*
- 54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*
- 55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*
- 56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*
- 57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*
- 58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*
- 59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*
- 60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*
- 61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*
- 62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*
- 63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*
- 64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*
- 65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*
- 66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*
- 67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*
- 68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*
- 69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*
- 70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*
- 71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*
- 72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*
- 73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*
- 74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*
- 75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3791.2	100.0	3792	17	US-09-351-224-10 Sequence 10, Appl
2	3791.2	100.0	3792	17	US-09-351-224E-10 Sequence 10, Appl
3	3791.2	100.0	3792	17	US-09-351-823-10 Sequence 10, Appl
4	3791.2	100.0	3792	26	US-09-677-488-10 Sequence 10, Appl
5	3791.2	100.0	3792	26	US-09-677-488A-10 Sequence 10, Appl
6	3791.2	100.0	3792	26	US-09-677-682-10 Sequence 10, Appl
7	3791.2	100.0	3792	26	US-09-677-682A-10 Sequence 10, Appl
8	3791.2	100.0	3792	26	US-09-677-682B-10 Sequence 10, Appl
9	3791.2	100.0	3792	33	US-09-882-694-10 Sequence 10, Appl
10	3791.2	100.0	3792	33	US-09-882-694A-10 Sequence 10, Appl
11	3454.8	91.1	3999	17	US-09-351-224-9 Sequence 9, Appl
12	3454.8	91.1	3999	17	US-09-351-224E-9 Sequence 9, Appl
13	3454.8	91.1	3999	17	US-09-351-823-9 Sequence 9, Appl
14	3454.8	91.1	3999	26	US-09-677-488-9 Sequence 9, Appl
15	3454.8	91.1	3999	26	US-09-677-488A-9 Sequence 9, Appl
16	3454.8	91.1	3999	26	US-09-677-682-9 Sequence 9, Appl
17	3454.8	91.1	3999	26	US-09-677-682A-9 Sequence 9, Appl
18	3454.8	91.1	3999	26	US-09-677-682B-9 Sequence 9, Appl
19	3454.8	91.1	3999	33	US-09-882-694-9 Sequence 9, Appl
20	3454.8	91.1	3999	33	US-09-882-694A-9 Sequence 9, Appl
21	723.6	19.1	3927	1	PCT-US98-27499-1 Sequence 1, Appl
22	722.2	19.0	3924	1	PCT-US98-27499-3 Sequence 3, Appl
23	487	12.8	4002	1	PCT-US98-27300-1 Sequence 1, Appl
24	487	12.8	4002	1	PCT-US98-27300-3 Sequence 3, Appl
25	451.4	11.9	3964	18	US-09-404-520-5444 Sequence 5444, Ap
26	450.6	11.9	4047	10	US-08-612-734-1 Sequence 1, Appl
27	367.4	9.7	4800	10	US-08-612-734-3 Sequence 3, Appl
28	362	9.5	3924	7	US-08-395-246A-1 Sequence 1, Appl
29	311.6	8.2	1818	18	US-09-417-507-4379 Sequence 4379, Ap
30	309	8.1	3942	23	US-09-614-150-18503 Sequence 18503, A
31	309	8.1	3942	56	US-60-173-464-15244 Sequence 15244, A



QY 1561 ggaagccttgagtgagtgacgacacaaagaaattgcaatcgacacgaagtatcatatcgat 1620  
DB 1561 ggaagccttgagtgagtgacgacacaaagaaattgcaatcgacacgaagtatcatatcgat 1620  
QY 1621 ccaagatcctgtatcatgaatgaagctaccagcgccttgaccggaagcggagaaagt 1680  
DB 1621 ccaagatcctgtatcatgaatgaagctaccagcgccttgaccggaagcggagaaagt 1680  
QY 1681 gtcagagagccttgacacgagtgccaaagaccgactactttgtgtcatgcccacaa 1740  
DB 1681 gtcagagagccttgacacgagtgccaaagaccgactactttgtgtcatgcccacaa 1740  
QY 1741 ctgagcactgtcaaaagtctggcaacatcgagctgacttccaggggaaatctgtcgag 1800  
DB 1741 ctgagcactgtcaaaagtctggcaacatcgagctgacttccaggggaaatctgtcgag 1800  
QY 1801 caagggcacacacacgaattgacgaattcggtctgactacgacgactggtgtgca 1860  
DB 1801 caagggcacacacacgaattgacgaattcggtctgactacgacgactggtgtgca 1860  
QY 1861 caggacctggggtgacgaacaaacaaagacatgagagaccctgcacgaacagcgacga 1920  
DB 1861 caggacctggggtgacgaacaaacaaagacatgagagaccctgcacgaacagcgacga 1920  
QY 1921 cgagaagctgctgtgaaacgacggcacttgagcgactcacacacactgcccactcaa 1980  
DB 1921 cgagaagctgctgtgaaacgacggcacttgagcgactcacacacactgcccactcaa 1980  
QY 1981 gctggagacactgagagagcgaagtgccggttcgggaacttgggtactgctcctctaaa 2040  
DB 1981 gctggagacactgagagagcgaagtgccggttcgggaacttgggtactgctcctctaaa 2040  
QY 2041 tgatcctaatcatgcttacgaacaaacaaatctctactggtgcttctgtgtgcaaca 2100  
DB 2041 tgatcctaatcatgcttacgaacaaacaaatctctactggtgcttctgtgtgcaaca 2100  
QY 2101 ataacggtctgatatgacggccacatctccagacagacgaccttctgttcagattg 2160  
DB 2101 ataacggtctgatatgacggccacatctccagacagacgaccttctgttcagattg 2160  
QY 2161 ctacgtctctgagtgagtgatgcacgacaggaacggcagacattttatattctg 2220  
DB 2161 ctacgtctctgagtgagtgatgcacgacaggaacggcagacattttatattctg 2220  
QY 2221 atgtctctgtgctctaggaatctagtagatatctacagatttgcagctgacatgc 2280  
DB 2221 atgtctctgtgctctaggaatctagtagatatctacagatttgcagctgacatgc 2280  
QY 2281 aacgttatttcaaaagtgtcacccatcgctatcaagcgcgaatgttccacgagtagtg 2340  
DB 2281 aacgttatttcaaaagtgtcacccatcgctatcaagcgcgaatgttccacgagtagtg 2340  
QY 2341 gatcaagacatgaactctcgacatcccgagcgaattctgtgtgctctcacatcgaa 2400  
DB 2341 gatcaagacatgaactctcgacatcccgagcgaattctgtgtgctctcacatcgaa 2400  
QY 2401 ctgtcagctctacccacgagtgcaagagtgctacacatagctatgctgataatctatc 2460  
DB 2401 ctgtcagctctacccacgagtgcaagagtgctacacatagctatgctgataatctatc 2460  
QY 2461 gttgtcgtcaacatcgctcgagcagtgctctacacatagctatgctgataatctg 2520  
DB 2461 gttgtcgtcaacatcgctcgagcagtgctctacacatagctatgctgataatctg 2520  
QY 2521 cctgt 2580  
DB 2521 cctgt 2580  
QY 2581 tctagagacgaagctagagacggaaactcgcaacttgcagaaagtgtgtgtgtgtgtgt 2640  
DB 2581 tctagagacgaagctagagacggaaactcgcaacttgcagaaagtgtgtgtgtgtgtgt 2640

QY 2641 aagcgaagcagctaccgcgacccgacgctctctctgactctcgaagcgaatgttct 2700  
DB 2641 aagcgaagcagctaccgcgacccgacgctctctctgactctcgaagcgaatgttct 2700  
QY 2701 caaacgactctcgaacatgttgagcaaggttctgtaagatcaacaaagcttttggtt 2760  
DB 2701 caaacgactctcgaacatgttgagcaaggttctgtaagatcaacaaagcttttggtt 2760  
QY 2761 ggaagaggttttggttttcaactgtcacagtcgagtgaggttttggctattgacctggga 2820  
DB 2761 ggaagaggttttggttttcaactgtcacagtcgagtgaggttttggctattgacctggga 2820  
QY 2821 tttgtattgagtcgagtcgaataattggcttcagtgagtcgacacacactcaatttatatca 2880  
DB 2821 tttgtattgagtcgagtcgaataattggcttcagtgagtcgacacacactcaatttatatca 2880  
QY 2881 tcttggtggcgttttgggttgcggtccaaagcagcagcccgatatttggcttactccag 2940  
DB 2881 tcttggtggcgttttgggttgcggtccaaagcagcagcccgatatttggcttactccag 2940  
QY 2941 agttttacaaagctcgtcggctcggaactatctctgctcggcgaacattgaacgg 3000  
DB 2941 agttttacaaagctcgtcggctcggaactatctctgctcggcgaacattgaacgg 3000  
QY 3001 accatccgcaaacgagaggaacaaagaaagcccgatgggtgatgacctgtcgac 3060  
DB 3001 accatccgcaaacgagaggaacaaagaaagcccgatgggtgatgacctgtcgac 3060  
QY 3061 ctgagagacattgaattcaggtatcgtaacggtatcggtcaggttctcccggggtt 3120  
DB 3061 ctgagagacattgaattcaggtatcgtaacggtatcggtcaggttctcccggggtt 3120  
QY 3121 tccatgacatcgagccaggaacattttagcttatgtggcgcttctggtgtggcaag 3180  
DB 3121 tccatgacatcgagccaggaacattttagcttatgtggcgcttctggtgtggcaag 3180  
QY 3181 tcaacgttgatcgcttttcggaagcattctacacccagcactcggccgaattctatt 3240  
DB 3181 tcaacgttgatcgcttttcggaagcattctacacccagcactcggccgaattctatt 3240  
QY 3241 gcaacgagaattatgcagaaatgtcgcgogcttgcacgcgacatgtctcttggct 3300  
DB 3241 gcaacgagaattatgcagaaatgtcgcgogcttgcacgcgacatgtctcttggct 3300  
QY 3301 caacgagaccacactttaccaagctccttcgagagaatgtgagcttggccctcgaa 3360  
DB 3301 caacgagaccacactttaccaagctccttcgagagaatgtgagcttggccctcgaa 3360  
QY 3361 gccaattatcagaagagctttgtcaagcgccttcccgcaagcgaatgtcttggatt 3420  
DB 3361 gccaattatcagaagagctttgtcaagcgccttcccgcaagcgaatgtcttggatt 3420  
QY 3421 ttgtatctcttaccagaagccttgaacgcttgcgctcgaacgaggtgagcttgc 3480  
DB 3421 ttgtatctcttaccagaagccttgaacgcttgcgctcgaacgaggtgagcttgc 3480  
QY 3481 tccggcggaacacagcagcggatfcgccatgcgaagcatttgcattgcgaatccaaagctg 3540  
DB 3481 tccggcggaacacagcagcggatfcgccatgcgaagcatttgcattgcgaatccaaagctg 3540  
QY 3541 ttgtacttgcgaagcagctcagccctcgacacgcaatcggaaactgtgttcaagct 3600  
DB 3541 ttgtacttgcgaagcagctcagccctcgacacgcaatcggaaactgtgttcaagct 3600  
QY 3601 gacctgagaggtccacgagcgaacatagcagtggtggcgaccgacttccact 3660  
DB 3601 gacctgagaggtccacgagcgaacatagcagtggtggcgaccgacttccact 3660  
QY 3661 attcggaattgt 3720  
DB 3661 attcggaattgt 3720  
QY 3721 cacgcggaactacacgactgagagaaagatatctacgagatgtgttttggcacaatttcta 3780

|||||  
Db 3721 cacgaggaactacaacgactgagaggaagatattacagagatgcttggcacaactttca 3780  
Qy 3781 gaccaagcatga 3792  
|||||  
Db 3781 gaccaagcatga 3792  
  
RESULT 2  
US-09-351-224E-10  
; Sequence 10, Application US/09351224E  
; GENERAL INFORMATION:  
; APPLICANT: Duvidk, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351,224E  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 3792  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA  
US-09-351-224E-10

Query Match 100.0%; Score 3791.2; DB 17; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 atggcagatgaatcgagaaactcgacaaacacgaagatggcagtgagtcctccacac 60  
|||||  
Db 1 atggcagatgaatcgagaaactcgacaaacacgaagatggcagtgagtcctccacac 60  
  
Qy 61 cctccccgaaaaagaaacgaagcagatttcagactatctcagactatctcagatat 120  
|||||  
Db 61 cctccccgaaaaagaaacgaagcagatttcagactatctcagactatctcagatat 120  
  
Qy 121 gccgacaaatacgaactggactctcaatgtcatcgcgctcatctgcgcatcggtaccggg 180  
|||||  
Db 121 gccgacaaatacgaactggactctcaatgtcatcgcgctcatctgcgcatcggtaccggg 180  
  
Qy 181 gcttccctctctgatgctgatcatcttcggtagcttcacaaacaaagttcaacaattac 240  
|||||  
Db 181 gcttccctctctgatgctgatcatcttcggtagcttcacaaacaaagttcaacaattac 240  
  
Qy 241 aattcggcgacggagtgactgaagcgttcaaggccgatgtgagatcatttcgctctg 300  
|||||  
Db 241 aattcggcgacggagtgactgaagcgttcaaggccgatgtgagatcatttcgctctg 300  
  
Qy 301 ttctgtacctcttatttgggaattgttcctcagctagcttttccacggtccattacc 360  
|||||  
Db 301 ttctgtacctcttatttgggaattgttcctcagctagcttttccacggtccattacc 360  
  
Qy 361 atttcagctatacgaacactcgaactcttcagacagatgttcttgatgaacacttg 420  
|||||  
Db 361 atttcagctatacgaacactcgaactcttcagacagatgttcttgatgaacacttg 420  
  
Qy 421 caagaggtctggcatttcacaaagcagagcaatggagcaatgccactcargtactacc 480  
|||||  
Db 421 caagaggtctggcatttcacaaagcagagcaatggagcaatgccactcargtactacc 480  
  
Qy 481 aatgcaacccgtatacaaacaggatttgcggagaaattggtttaccgtcaggacatt 540  
|||||

Db 481 aatggcaaccgtatatacaaaacagggtattgcccagagaaattggtctttaccgtgcaggcactt 540  
Qy 541 tcaatgtcttttttgcatttgggttgcttggcgtctcagtggaagctagcttttaac 600  
|||||  
Db 541 tcaatgtcttttttgcatttgggttgcttggcgtctcagtggaagctagcttttaac 600  
|||||  
Qy 601 accatgctcgtcatccctgcctcatttctcgttcacccggcactcgtcagagaattgagtc 660  
|||||  
Db 601 accatgctcgtcatccctgcctcatttctcgttcacccggcactcgtcagagaattgagtc 660  
|||||  
Qy 661 gctcaggaggccaggatcaccaggatctactacgcgcgcgtctcctcgcaagaagtc 720  
|||||  
Db 661 gctcaggaggccaggatcaccaggatctactacgcgcgcgtctcctcgcaagaagtc 720  
|||||  
Qy 721 ttatcatccatcccgacagtcctcatttctacgccagagaaaaatggttcgaaaaatat 780  
|||||  
Db 721 ttatcatccatcccgacagtcctcatttctacgccagagaaaaatggttcgaaaaatat 780  
|||||  
Qy 781 gatgtcttttgcaggaagcacaacagaaagggaagaaatcgcaaaataatggagtc 840  
|||||  
Db 781 gatgtcttttgcaggaagcacaacagaaagggaagaaatcgcaaaataatggagtc 840  
|||||  
Qy 841 ttgttcaactgagtagtcttctgcatcttctacgtctgactcgcactggttttggaaaggt 900  
|||||  
Db 841 ttgttcaactgagtagtcttctgcatcttctacgtctgactcgcactggttttggaaaggt 900  
|||||  
Qy 901 ttctgcgtatcagaaatggcggaggttgcgcagcttggcgaagctttactggttctct 960  
|||||  
Db 901 ttctgcgtatcagaaatggcggaggttgcgcagcttggcgaagctttactggttctct 960  
|||||  
Qy 961 tccgtcaccttagcagcgcagctccatctcaatgcttgcgccttcaggttcaggtcttac 1020  
|||||  
Db 961 tccgtcaccttagcagcgcagctccatctcaatgcttgcgccttcaggttcaggtcttac 1020  
|||||  
Qy 1021 caacgcagcatcttcggtccgaattattcagatcatctgacaaacacacacacagctgcac 1080  
|||||  
Db 1021 caacgcagcatcttcggtccgaattattcagatcatctgacaaacacacacacagctgcac 1080  
|||||  
Qy 1081 cctctcgaccttctggaagcagcagagggctgcttagtcaaatggagatcacaacac 1140  
|||||  
Db 1081 cctctcgaccttctggaagcagcagagggctgcttagtcaaatggagatcacaacac 1140  
|||||  
Qy 1141 ctggcatttgcctccctccgccatctgccaaagtaetctcagatttcaacttgaca 1200  
|||||  
Db 1141 ctggcatttgcctccctccgccatctgccaaagtaetctcagatttcaacttgaca 1200  
|||||  
Qy 1201 attcagctggcagagacagcgcctcgtcgtgcatcaggtagcgcgcaaaagcacaaatg 1260  
|||||  
Db 1201 attcagctggcagagacagcgcctcgtcgtgcatcaggtagcgcgcaaaagcacaaatg 1260  
|||||  
Qy 1261 gtcggttacttgacggtggtatctgccagttcggggagagatattacttgatgggttg 1320  
|||||  
Db 1261 gtcggttacttgacggtggtatctgccagttcggggagagatattacttgatgggttg 1320  
|||||  
Qy 1321 gaactgggaacatacaatgtgaaatgctgagaaagccgcatctcgcctcgttcaacagaa 1380  
|||||  
Db 1321 gaactgggaacatacaatgtgaaatgctgagaaagccgcatctcgcctcgttcaacagaa 1380  
|||||  
Qy 1381 cctgtgttcttctggtggcacaatcttccagaaacttgcacacggttttcaatggatgagaa 1440  
|||||  
Db 1381 cctgtgttcttctggtggcacaatcttccagaaacttgcacacggttttcaatggatgagaa 1440  
|||||  
Qy 1441 cgaatctgcctcgcgaaaaaacaatggagcttctgcaaaagcttcgaagcagcaaat 1500  
|||||  
Db 1441 cgaatctgcctcgcgaaaaaacaatggagcttctgcaaaagcttcgaagcagcaaat 1500  
|||||  
Qy 1501 ggcgacgtgttcattatgagcttccgaacggtttatgagactgaagtggcgagcgc 1560  
|||||  
Db 1501 ggcgacgtgttcattatgagcttccgaacggtttatgagactgaagtggcgagcgc 1560  
|||||  
Qy 1561 ggaaccttgagtgaggtcaacgacgaatgcaatcgacacgaagtatcatatcatcgat 1620  
|||||  
Db 1561 ggaaccttgagtgaggtcaacgacgaatgcaatcgacacgaagtatcatatcatcgat 1620  
|||||

QY 1621 cccaagatcctgttactcgtatgaagctaccagcgcccttgaccggaagcgagaaagtg 1680  
Db 1621 cccaagatcctgttactcgtatgaagctaccagcgcccttgaccggaagcgagaaagtg 1680  
QY 1661 gtccagagagccctgaaccgagtggtccaaagaccgactacttggctcatgcccacaaa 1740  
Db 1661 gtccagagagccctgaaccgagtggtccaaagaccgactacttggctcatgcccacaaa 1740  
QY 1741 ctatgccactgtcaaaagtgtctggcaacatcgcagtcatttcccaggggaaatcgtcgag 1800  
Db 1741 ctatgccactgtcaaaagtgtctggcaacatcgcagtcatttcccaggggaaatcgtcgag 1800  
QY 1801 caagggcacacacacgaattgacgaattcggctgtcattacgcgcgactggtcgctgca 1860  
Db 1801 caagggcacacacacgaattgacgaattcggctgtcattacgcgcgactggtcgctgca 1860  
QY 1861 caggacctcgggtgacgacacaaagaaacatgagaagaccctgcagaaagagcgaca 1920  
Db 1861 caggacctcgggtgacgacacaaagaaacatgagaagaccctgcagaaagagcgaca 1920  
QY 1921 cgagaagctgctggtgaacgaccggaacttgagcgcaactcacaccactgcacacatcaca 1980  
Db 1921 cgagaagctgctggtgaacgaccggaacttgagcgcaactcacaccactgcacacatcaca 1980  
QY 1981 gctgagagacctgagaaagcggaagtgcggtcggaacttgggctactcgtcctctaaa 2040  
Db 1981 gctgagagacctgagaaagcggaagtgcggtcggaacttgggctactcgtcctctaaa 2040  
QY 2041 tgcactcctaactcatgttctacgaaacaaataatctctactgggtcttctgttctaaaca 2100  
Db 2041 tgcactcctaactcatgttctacgaaacaaataatctctactgggtcttctgttctaaaca 2100  
QY 2101 ataacggttctgatgtcgcggccacatttccagacaaagccctttgtttctcgagattg 2160  
Db 2101 ataacggttctgatgtcgcggccacatttccagacaaagccctttgtttctcgagattg 2160  
QY 2161 cteactgtctcagttgagtgatgtaatcggcacaggaacggcgagacttttatattctg 2220  
Db 2161 cteactgtctcagttgagtgatgtaatcggcacaggaacggcgagacttttatattctg 2220  
QY 2221 atgttcttctgtcgtctctagaaatctagtaggatatattcaagattcggctgagacatgc 2280  
Db 2221 atgttcttctgtcgtctctagaaatctagtaggatatattcaagattcggctgagacatgc 2280  
QY 2281 aacgttatttcaaaagtgttccaccatcgtctatcaagcgcgaattgttccaaacgagtactg 2340  
Db 2281 aacgttatttcaaaagtgttccaccatcgtctatcaagcgcgaattgttccaaacgagtactg 2340  
QY 2341 gatcaagacatcgaactctcgcacatcccgagagcaaatcttggtgctctcacatcgcaa 2400  
Db 2341 gatcaagacatcgaactctcgcacatcccgagagcaaatcttggtgctctcacatcgcaa 2400  
QY 2401 ctgtcagctctacccacgagttgcaagagttgatcagcaaaatttctcatttatctc 2460  
Db 2401 ctgtcagctctacccacgagttgcaagagttgatcagcaaaatttctcatttatctc 2460  
QY 2461 gttgtcgggtcaacatcgtctcagagcagtgctctacacatagcttatgatggaacatggg 2520  
Db 2461 gttgtcgggtcaacatcgtctcagagcagtgctctacacatagcttatgatggaacatggg 2520  
QY 2521 ccttggttggtgtgttgggtgacattccacccctgtcttgggtggtcactcgaagattcg 2580  
Db 2521 ccttggttggtgtgttgggtgacattccacccctgtcttgggtggtcactcgaagattcg 2580  
QY 2581 tctagagacgaagctagagcggaactcggcaacttgcagaaagtgcgtgggcttgc 2640  
Db 2581 tctagagacgaagctagagcggaactcggcaacttgcagaaagtgcgtgggcttgc 2640  
QY 2641 aagcgagagcagttaccgcgaatccgaccctctcatcttcttgactctcgaagccatgttct 2700  
Db 2641 aagcgagagcagttaccgcgaatccgaccctctcatcttcttgactctcgaagccatgttct 2700

QY 2701 ccaacagctactcggacatgttgagcaagggtcttgyctaagatcatccaagcgttttggtt 2760  
Db 2701 ccaacagctactcggacatgttgagcaagggtcttgyctaagatcatccaagcgttttggtt 2760  
QY 2761 ggaacgatgtttggttttctactgttcacagtcgatggagtttttggctattgcccctggga 2820  
Db 2761 ggaacgatgtttggttttctactgttcacagtcgatggagtttttggctattgcccctggga 2820  
QY 2821 ttctgtatgtcagtcgatataattgcttcaagtgagtgacgacacaaactcaattttatata 2880  
Db 2821 ttctgtatgtcagtcgatataattgcttcaagtgagtgacgacacaaactcaattttatata 2880  
QY 2881 tcttcggtggggtttgtttgcccgttccaagcagcagccagttatttgccttactccag 2940  
Db 2881 tcttcggtggggtttgtttgcccgttccaagcagcagccagttatttgccttactccag 2940  
QY 2941 agttttaccaaggtctcggtcggtcgcaactatatactctgctcgagacattggaagcg 3000  
Db 2941 agttttaccaaggtctcggtcggtcgcaactatatactctgctcgagacattggaagcg 3000  
QY 3001 accatcccgaaacggaggaacaaagaaagggcccagtggtggtggtggtgcctgtcgac 3060  
Db 3001 accatcccgaaacggaggaacaaagaaagggcccagtggtggtggtggtgcctgtcgac 3060  
QY 3061 ctccgagacattgaattcaggtatcgtcaacgtgattcggctcgagttctccgcgsggtt 3120  
Db 3061 ctccgagacattgaattcaggtatcgtcaacgtgattcggctcgagttctccgcgsggtt 3120  
QY 3121 tccatgacaaatccgacgacgagacaatttctgagcttatgtgggcgtctctaggctgtggcaag 3180  
Db 3121 tccatgacaaatccgacgacgagacaatttctgagcttatgtgggcgtctctaggctgtggcaag 3180  
QY 3181 tcaacgttgatcgtcttgcggaacgatttcaagaccgacctcgggcccgaatttccatt 3240  
Db 3181 tcaacgttgatcgtcttgcggaacgatttcaagaccgacctcgggcccgaatttccatt 3240  
QY 3241 gcacacgagaattatgcagaaatgtccgcgcttgaacgcgggacatatgtcttttggtc 3300  
Db 3241 gcacacgagaattatgcagaaatgtccgcgcttgaacgcgggacatatgtcttttggtc 3300  
QY 3301 caaacggaacccacacactttaccaaaggtcgttcgcgagaaatgtgagctgtggccctcgaa 3360  
Db 3301 caaacggaacccacacactttaccaaaggtcgttcgcgagaaatgtgagctgtggccctcgaa 3360  
QY 3361 gccgaattatcagaagagcttgcgaaggacgaccttcccgaaggccaatgtcttgatt 3420  
Db 3361 gccgaattatcagaagagcttgcgaaggacgaccttcccgaaggccaatgtcttgatt 3420  
QY 3421 ttgtcatctctttaccagaagcttgaacgcttgcgctcaacgagggatgcagttc 3480  
Db 3421 ttgtcatctctttaccagaagcttgaacgcttgcgctcaacgagggatgcagttc 3480  
QY 3481 tccggggcaacgacagcgatcgcctatcgaagagcattgtatcgaaatccaaagctg 3540  
Db 3481 tccggggcaacgacagcgatcgcctatcgaagagcattgtatcgaaatccaaagctg 3540  
QY 3541 ttgtcacttgcagaaagcagctcagccctcgacacgcaatcggaaactcgttctgaagct 3600  
Db 3541 ttgtcacttgcagaaagcagctcagccctcgacacgcaatcggaaactcgttctgaagct 3600  
QY 3601 gccctcgatgaggtatccagcgacgacgaataagacagtgagcagcacttccact 3660  
Db 3601 gccctcgatgaggtatccagcgacgacgaataagacagtgagcagcacttccact 3660  
QY 3661 attcggaaatgttgattttttttgttttgcgaacgggagaaatccccaacgggcaact 3720  
Db 3661 attcggaaatgttgattttttttgttttgcgaacgggagaaatccccaacgggcaact 3720  
QY 3721 caccggaactacacgactgagaggaagatatcagagatgtgtttggcacaactttta 3780  
Db 3721 caccggaactacacgactgagaggaagatatcagagatgtgtttggcacaactttta 3780  
QY 3781 gaccgaagcatga 3792

Db 3781 gaccaagcatga 3792

|||||

## RESULT 3

US-09-351-823-10

: Sequence 10, Application US/09351823

: GENERAL INFORMATION:

: APPLICANT: Duwick, Jon

: TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification

: FILE REFERENCE: 5718-56

: CURRENT APPLICATION NUMBER: US/09/351,823

: CURRENT FILING DATE: 1999-07-12

: NUMBER OF SEQ ID NOS: 11

: SOFTWARE: PatentIn Ver. 2.0

: SEQ ID NO 10

: LENGTH: 3792

: TYPE: DNA

: ORGANISM: *Exophiala spinifera*

: FEATURE:

: OTHER INFORMATION: p-glycoprotein, coding sequence

US-09-351-823-10

Query Match 100.0%; Score 3791.2; DB 17; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1	atggcagatgaatcgagaaacctcgacccaaacgaagatggcagtgagctctctcacac	60
Db 1	atggcagatgaatcgagaaacctcgacccaaacgaagatggcagtgagctctcacac	60
Qy 61	ctctcccccagaaagaaacgaagcagatttcagactatctacgaatttcagatat	120
Db 61	ctctcccccagaaagaaacgaagcagatttcagactatctacgaatttcagatat	120
Qy 121	gcgacaaatcacgactggagctctcaatgctacgctgcctatctgcgcacgttcgag	180
Db 121	gcgacaaatcacgactggagctctcaatgctacgctgcctatctgcgcacgttcgag	180
Qy 181	gtctccctctctgtatgcgactatctcgttagcttccaaacaaagttcaacaattac	240
Db 181	gtctccctctctgtatgcgactatctcgttagcttccaaacaaagttcaacaattac	240
Qy 241	aattcggcgacggagctgaagcgtttcaaggccgagatggatcatcttgcctctg	300
Db 241	aattcggcgacggagctgaagcgtttcaaggccgagatggatcatcttgcctctg	300
Qy 301	ttcgtctacctctttattgggaagtttgctcactgacgttttccacggctgcaattacc	360
Db 301	ttcgtctacctctttattgggaagtttgctcactgacgttttccacggctgcaattacc	360
Qy 361	atttcagctatagcaaccactgcaactcttcgacagtggttccctgaatgcaacttcg	420
Db 361	atttcagctatagcaaccactgcaactcttcgacagtggttccctgaatgcaacttcg	420
Qy 421	caagagctctggatttcgacaaagagcaaatggagcaatcgccactcargtcaattacc	480
Db 421	caagagctctggatttcgacaaagagcaaatggagcaatcgccactcargtcaattacc	480
Qy 481	aatggcaacgtatatacaacagatttggccgagaaattgggtcttaccgtcagcgactt	540
Db 481	aatggcaacgtatatacaacagatttggccgagaaattgggtcttaccgtcagcgactt	540
Qy 541	tcaatgtctcttctgcatctggtcttggcgtctcagtgaggaagctagcttaacc	600
Db 541	tcaatgtctcttctgcatctggtcttggcgtctcagtgaggaagctagcttaacc	600
Qy 601	accatgtccgtctacccctgcatcttctcgttgcacccgcatctcgtatagcaattgagc	660
Db 601	accatgtccgtctacccctgcatcttctcgttgcacccgcatctcgtatagcaattgagc	660
Qy 661	gtccaggggcccagatccacggatctactcaacgcgcgctgctcctcgagagaagtc	720

Db 661	gtccaggggcccagatccacggatctactcaacgcgcgctgctcctcgagagaagtc	720
Qy 721	ttatcatcatccggagacagtcacgtcttcttaagccacagaagaaatggtcgaaaaatat	780
Db 721	ttatcatcatccggagacagtcacgtcttcttaagccacagaagaaatggtcgaaaaatat	780
Qy 781	gatgtctttttgcagcagcacacaaagagggagagaaatcgccaaataatgsgsgtc	840
Db 781	gatgtctttttgcagcagcacacaaagagggagagaaatcgccaaataatgsgsgtc	840
Qy 841	ttgttctcaactgagtagctcttttgcatcttaagctgcctatcgcaactgcccctttgagaagtc	900
Db 841	ttgttctcaacgtagctcttttgcatcttaagctgcctatcgcaactgcccctttgagaagtc	900
Qy 901	tttcgcatgtatcagaatggcaggttggcgaactgtggcgaagctcttactgttgcctt	960
Db 901	tttcgcatgtatcagaatggcaggttggcgaactgtggcgaagctcttactgttgcctt	960
Qy 961	tcctgcaactttagcagccacagtcacatctcaatgcttgcgcttccaggttcagtcgtttac	1020
Db 961	tcctgcaactttagcagccacagtcacatctcaatgcttgcgcttccaggttcagtcgtttac	1020
Qy 1021	caacgcgcacactctcggctcgcgaattattcagtatcatctgacaaaaccacgagctcgac	1080
Db 1021	caacgcgcacactctcggctcgcgaattattcagtatcatctgacaaaaccacgagctcgac	1080
Qy 1081	ctctctgacaccttctgaaagcagcagaggtgctgctaggttcaaatctgagatcccaaac	1140
Db 1081	ctctctgacaccttctgaaagcagcagaggtgctgctaggttcaaatctgagatcccaaac	1140
Qy 1141	ctggcatttgcctaccctccgaccatctgcccgaagtaacttcgagatttcaacttgaca	1200
Db 1141	ctggcatttgcctaccctccgaccatctgcccgaagtaacttcgagatttcaacttgaca	1200
Qy 1201	attccagctggcgaagcagcagcgcctcgtcgttgcatacaggttagcggcaaaagcacaatc	1260
Db 1201	attccagctggcgaagcagcagcgcctcgtcgttgcatacaggttagcggcaaaagcacaatc	1260
Qy 1261	gtcggcttacttgaaacggttggtatctgccagtttcggggagagatattacttgatgggttc	1320
Db 1261	gtcggcttacttgaaacggttggtatctgccagtttcggggagagatattacttgatgggttc	1320
Qy 1321	gaactggagcaacaatgaaatggaatggaagcgaattccgctcttcaacagga	1380
Db 1321	gaactggagcaacaatgaaatggaatggaagcgaattccgctcttcaacagga	1380
Qy 1381	ctcgtgttcttctggaacaaattctccagaacattgccaacggttccatgatgagcaa	1440
Db 1381	ctcgtgttcttctggaacaaattctccagaacattgccaacggttccatgatgagcaa	1440
Qy 1441	cgagatctgcctcgcgaaaaaaatggagcttctgcaaaaagcttgcgaagccagcaat	1500
Db 1441	cgagatctgcctcgcgaaaaaaatggagcttctgcaaaaagcttgcgaagccagcaat	1500
Qy 1501	ggcgagctgttcaataatgagcttccgaacggttatagagactgaagcttggcgagcgacc	1560
Db 1501	ggcgagctgttcaataatgagcttccgaacggttatagagactgaagcttggcgagcgacc	1560
Qy 1561	ggagcttctgagtgaggttcaacgcaacaatgcaatcgcaatcgcaagctatcatatcgat	1620
Db 1561	ggagcttctgagtgaggttcaacgcaacaatgcaatcgcaatcgcaagctatcatatcgat	1620
Qy 1621	cccaagatctcctgttactcgaatgagcttaccagcgccttgaccgagcgaggaagtcg	1680
Db 1621	cccaagatctcctgttactcgaatgagcttaccagcgccttgaccgagcgaggaagtcg	1680
Qy 1681	gtccaggggaccttgacccagatgtccaaagacgcgaactactttggttacttgcacaaa	1740
Db 1681	gtccaggggaccttgacccagatgtccaaagacgcgaactactttggttacttgcacaaa	1740
Qy 1741	ctagcagctgtccaaaagtgtgtggcaacatcgcatcttcccgaggggaaaaatcgtcgag	1800



Db	1741	ctagccactgtctcaaaagtgtctggoaacatcgacgtlcatcttcccagggggaaaaatcgtctgag	1800
Qy	1801	caaggcacacaccacgaattgatcgaattcgctgtcattacgcccgaactgggtgcgtgca	1860
Db	1801	caaggcacacaccacgaattgatcgaattcgctgtcattacgcccgaactgggtgcgtgca	1860
Qy	1861	caggacctcgggctgcagaaacaaagaacatgagaagacctgcacgaaaaggcagca	1920
Db	1861	caggacctcgggctgcagaaacaaagaacatgagaagacctgcacgaaaaggcagca	1920
Qy	1921	cgaagaagctgtcgttgaaacgaccggcacttgagcgcactcacaccaactgcacatctcaa	1980
Db	1921	cgaagaagctgtcgttgaaacgaccggcacttgagcgcactcacaccaactgcacatctcaa	1980
Qy	1981	gctcgagaagcctggagaagcgaaggtccggtcgaggacttgggctactcgctccctaaaa	2040
Db	1981	gctcgagaagcctggagaagcgaaggtccggtcgaggacttgggctactcgctccctaaaa	2040
Qy	2041	tgcactctaatacatgtcttcagaaacaaaaatctctactgtgtctcttgttgtcaaca	2100
Db	2041	tgcactctaatacatgtcttcagaaacaaaaatctctactgtgtctcttgttgtcaaca	2100
Qy	2101	ataacggttctcatatgcgcgcgcacatttcagagacaagccctttgttttcgagattg	2160
Db	2101	ataacggttctcatatgcgcgcgcacatttcagagacaagccctttgttttcgagattg	2160
Qy	2161	ctcaatgtctctcgagttgagtggtctcgtgcacaggaacgggcagacttttatattctg	2220
Db	2161	ctcaatgtctctcgagttgagtggtctcgtgcacaggaacgggcagacttttatattctg	2220
Qy	2221	atgttttttgtcgtggctctagaaaatctagttagatatatttcacgatttgctggacatgc	2280
Db	2221	atgttttttgtcgtggctctagaaaatctagttagatatatttcacgatttgctggacatgc	2280
Qy	2281	aacgttatttccaaagtgtcaaccatgcgtatcaaacgcgcaatgttccaaacgagtaactg	2340
Db	2281	aacgttatttccaaagtgtcaaccatgcgtatcaaacgcgcaatgttccaaacgagtaactg	2340
Qy	2341	gatacagacatcgaaactctcgacatcccgcgagcaaaattctgtgctctcaacatcgcaa	2400
Db	2341	gatacagacatcgaaactctcgacatcccgcgagcaaaattctgtgctctcaacatcgcaa	2400
Qy	2401	ctgtcaactctccaaacgagttgcaagagttgatacaagcaaatcttcaatttatctc	2460
Db	2401	ctgtcaactctccaaacgagttgcaagagttgatacaagcaaatcttcaatttatctc	2460
Qy	2461	gttgtcgggtccaacatcgtctcgagcagtgctctcaccactgcctatggatggaaactggg	2520
Db	2461	gttgtcgggtccaacatcgtctcgagcagtgctctcaccactgcctatggatggaaactggg	2520
Qy	2521	ccctgttggtgtgtttgtgtgcaacttccaaacctgcttttggtgggtacactcgaattctg	2580
Db	2521	ccctgt	2580
Qy	2581	tctagagacgaagctagaaacgggaactcggcaaaactttgcagaaagtgtcgggcttgc	2640
Db	2581	tctagagacgaagctagaaacgggaactcggcaaaactttgcagaaagtgtcgggcttgc	2640
Qy	2641	aagcgaagcagttaccggatccggaacgtctcatctttgaactctcgaagccatgtctct	2700
Db	2641	aagcgaagcagttaccggatccggaacgtctcatctttgaactctcgaagccatgtctct	2700
Qy	2701	ccaaacgactctcggaactgtgagcaaggtctgtcaagatcccaaaagcttttgggtt	2760
Db	2701	ccaaacgactctcggaactgtgagcaaggtctgtcaagatcccaaaagcttttgggtt	2760
Qy	2761	ggacgagtgttttgggttttcaactgtccacagtcgattggagtttttttggcatttgccttggga	2820
Db	2761	ggacgagtgttttgggttttcaactgtccacagtcgattggagtttttttggcatttgccttggga	2820
Qy	2821	ttttgtattgcagtcogataatttgggttccaggttgagtgacgacacaacccaattttatatca	2880
Db	2821	ttttgtattgcagtcogataatttgggttccaggttgagtgacgacacaacccaattttatatca	2880

RESULTS

RESULT 4  
US-09-677-488-10

US-09-677-488-10  
: Sequence 10, Application US/09677488; sequence 10, applica  
: GENERAL INFORMATION:  
: GENERAL INFORMATION:

GENERAL INFORMATION.

APPLICANT: Duvick, Jon

APPLICANT: DUVICK, DON  
TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification

FILE REFERENCE: 35718/204100 (5718-111A)

```

; CURRENT APPLICATION NUMBER: US/09/677,488
; CURRENT FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 09/351,234
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3792
; TYPE: DNA
; ORGANISM: Exophiala sp.
; FEATURE:
; OTHER INFORMATION: p-glycoprotein, fully
US-09-677-488-10

```

Query Match	100.0%	Score 3791.2	DB 26	Length 3792
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 3792	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	atggcagatgaatcggagagaaacccctgcaccaaaccagaatgacgagtcgctctcacac	60	
Db	1	atggcagatgaatcggagagaaacccctgcaccaaaccagaatgacgagtcgctctcacac	60	
QY	61	ctctcccccagaaaaagaaacccgagggcaggtattttcagactatctacgaatcttcagatat	120	
Db	61	ctctcccccagaaaaagaaacccgagggcaggtattttcagactatctacgaatcttcagatat	120	
QY	121	gcgcacaaatacgaatcgactctcaatgtcatcgcgcctcatctcgccatcggagbcggg	180	
Db	121	gcgcacaaatacgaatcgactctcaatgtcatcgcgcctcatctcgccatcggagtcggg	180	
QY	181	gcttcctctcttgatctgcatcatcttcggtagctctccaaacgaagtcacaattac	240	
Db	181	gcttcctctcttgatctgcatcatcttcggtagctctccaaacgaagtcacaattac	240	
QY	241	aattcggcgacgggagtcctgaagcgttcaaggccgatggcatcttcgttcctgtgg	300	
Db	241	aattcggcgacgggagtcctgaagcgttcaaggccgatggcatcttcgttcctgtgg	300	
QY	301	ttcgtctactctttattgggaagtcttcctcaagtaagctttccacggctgcattacc	360	
Db	301	ttcgtctactctttattgggaagtcttcctcaagtaagctttccacggctgcattacc	360	
QY	361	atttcagctatcacgaacactcgaactcttcgacgaggttccttgaatgcacctggg	420	
Db	361	atttcagctatcacgaacactcgaactcttcgacgaggttccttgaatgcacctggg	420	
QY	421	caagaggctggcatcttcgacagcagagcaatggagcaatcgccactcargtcaatacc	480	
Db	421	caagaggctggcatcttcgacagcagagcaatggagcaatcgccactcargtcaatacc	480	
QY	481	aatggcaacccgtatacaaacagggtattgcgcagagaaatggctttaccgtgcaggcaact	540	
Db	481	aatggcaacccgtatacaaacagggtattgcgcagagaaatggctttaccgtgcaggcaact	540	
QY	541	tcaatgtctctttctgcatttgttgttcgcttcggcgtctcaagtggagtagctttaatc	600	
Db	541	tcaatgtctctttctgcatttgttgttcgcttcggcgtctcaagtggagtagctttaatc	600	
QY	601	accatgtcgtcatctctgcctatttcttggttcacccggcactcgcagtagcaattgatgcc	660	
Db	601	accatgtcgtcatctctgcctatttcttggttcacccggcactcgcagtagcaattgatgcc	660	
QY	661	gctcaggagggcaggatcacacaggatctactcacgcgcgctctctccagagaagtc	720	
Db	661	gctcaggagggcaggatcacacaggatctactcacgcgcgctctctccagagaagtc	720	
QY	721	ttatcatccatccggagacgttccatgtctttctatagcccaagaagaaatgtctgaaaaatat	780	
Db	721	ttatcatccatccggagacgttccatgtctttctatagcccaagaagaaatgtctgaaaaatat	780	
QY	781	gatgtctcttttcgacgaagcacacacgaagggaagaagaaatcgcaccaataatggsgtc	840	

QY 1921 c9agaagctgctggtgaaacgaccgcaacttgagcgcactcaacacactgcacactctcaa 1980  
Db 1921 c9agaagctgctggtgaaacgaccgcaacttgagcgcactcaacacactgcacactctcaa 1980  
QY 1981 gctggagacgtggaagcgggaagtgccggtcgggacttgggctactcgcctctaaa 2040  
Db 1981 gctggagacgtggaagcgggaagtgccggtcgggacttgggctactcgcctctaaa 2040  
QY 2041 tgcactcctaataatgtttctacgaaacaaataatctctactgtgtctctgttgaaca 2100  
Db 2041 tgcactcctaataatgtttctacgaaacaaataatctctactgtgtctctgttgaaca 2100  
QY 2101 ataacggtctgatatgcggccacatttccagagacaagccctttgtttcagagt 2160  
Db 2101 ataacggtctgatatgcggccacatttccagagacaagccctttgtttcagagt 2160  
QY 2161 ctcaactgtctcagttgagttggtcatcgccacaggaacgggcagacttttatcttg 2220  
Db 2161 ctcaactgtctcagttgagttggtcatcgccacaggaacgggcagacttttatcttg 2220  
QY 2221 atgtttctgtcgtggtctagaaatctagtaggatatattcaagattggctgacatgc 2280  
Db 2221 atgtttctgtcgtggtctagaaatctagtaggatatattcaagattggctgacatgc 2280  
QY 2281 aacgttattccaaagtgtccacatcgtctatcaagcgcgaatgttccaaacgagtactg 2340  
Db 2281 aacgttattccaaagtgtccacatcgtctatcaagcgcgaatgttccaaacgagtactg 2340  
QY 2341 gatacagacactcgaaactcgtcagatcccgaggagaaatttctgtgctctcaacatcaa 2400  
Db 2341 gatacagacactcgaaactcgtcagatcccgaggagaaatttctgtgctctcaacatcaa 2400  
QY 2401 ctgtcagacttaccacgcagttgcaagagttgatatacagaataattttctcatattatc 2460  
Db 2401 ctgtcagacttaccacgcagttgcaagagttgatatacagaataattttctcatattatc 2460  
QY 2461 qttcgtcgttcacacatcgtctcagcagtgctctacactagctatgatggaacggg 2520  
Db 2461 qttcgtcgttcacacatcgtctcagcagtgctctacactagctatgatggaacggg 2520  
QY 2521 cctgtgtgtgtgtgtgtgtgcaacttccacccctgttttggctgctacactcagaattcg 2580  
Db 2521 cctgtgtgtgtgtgtgtgtgcaacttccacccctgttttggctgctacactcagaattcg 2580  
QY 2581 tctagagacgaagctagaaacgggaactcggcaacttgcagaaagtgcgggcttcg 2640  
Db 2581 tctagagacgaagctagaaacgggaactcggcaacttgcagaaagtgcgggcttcg 2640  
QY 2641 aagcgaagcagttaccgcgactccggacogtctcactcttttgactctcgaagccatgttct 2700  
Db 2641 aagcgaagcagttaccgcgactccggacogtctcactcttttgactctcgaagccatgttct 2700  
QY 2701 ccaacagtaactcggacatgttgagcaagttcttctaaagataccaagcttttggttt 2760  
Db 2701 ccaacagtaactcggacatgttgagcaagttcttctaaagataccaagcttttggttt 2760  
QY 2761 ggaagatgtttgttttcaactgtcacagtgatgagatgttttgcatttgcctggga 2820  
Db 2761 ggaagatgtttgttttcaactgtcacagtgatgagatgttttgcatttgcctggga 2820  
QY 2821 ttttgtattgcagtcgaataattggcttcaggtgagtaacgacacaaactcaattttatca 2880  
Db 2821 ttttgtattgcagtcgaataattggcttcaggtgagtaacgacacaaactcaattttatca 2880  
QY 2881 tcttcgtggcggttttggcttcgggtccgaagcagcagcagtttggcttactccacg 2940  
Db 2881 tcttcgtggcggttttggcttcgggtccgaagcagcagcagtttggcttactccacg 2940  
QY 2941 agttttaccaggtcgtggtcgtgcgaactatatactctggtcgtgagacattgaagcgg 3000  
Db 2941 agttttaccaggtcgtggtcgtgcgaactatatactctggtcgtgagacattgaagcgg 3000

QY 3001 accatccgcgaacgagagagacacaaaaagggcccggtgggtgagtcctcctgcgac 3060  
Db 3001 accatccgcgcgaacgagagagacacaaaaagggcccggtgggtgagtcctcctgcgac 3060  
QY 3061 ctcgaggacattgaattcaagtgatcgttcaacgtgattcggctcagtttcccggggtt 3120  
Db 3061 ctcgaggacattgaattcaagtgatcgttcaacgtgattcggctcagtttcccggggtt 3120  
QY 3121 tccatgacaatcgagccagggacaattttagcttattatggtggtccttctgctgtggcaag 3180  
Db 3121 tccatgacaatcgagccagggacaattttagcttattatggtggtccttctgctgtggcaag 3180  
QY 3181 tcaacgttgatcgtttgtcgaagcatttcaacccgacctcggccgaatttcatitt 3240  
Db 3181 tcaacgttgatcgtttgtcgaagcatttcaacccgacctcggccgaatttcatitt 3240  
QY 3241 gcacacgagaattatgcgagaatgtcgccgcttgcacgcggccatattgttttggtc 3300  
Db 3241 gcacacgagaattatgcgagaatgtcgccgcttgcacgcggccatattgttttggtc 3300  
QY 3301 caacaggaacccacactttaccagggtccgttcgcgagaaatgcagcttggccctcgaa 3360  
Db 3301 caacaggaacccacactttaccagggtccgttcgcgagaaatgcagcttggccctcgaa 3360  
QY 3361 ggcgaattatcagaagagctttgttcaaggacgccttcccgcaaggccaaatgcttggatt 3420  
Db 3361 ggcgaattatcagaagagctttgttcaaggacgccttcccgcaaggccaaatgcttggatt 3420  
QY 3421 ttgtcatctctttaccagaagcctttgaaacgcttgcggtcgaagggatgcagttc 3480  
Db 3421 ttgtcatctctttaccagaagcctttgaaacgcttgcggtcgaagggatgcagttc 3480  
QY 3481 tccgcggaacacgacgagctgcacatgcacgcgaatcggaacgtctgttgcagct 3540  
Db 3481 tccgcggaacacgacgagctgcacatgcacgcgaatcggaacgtctgttgcagct 3540  
QY 3541 ttgtacttgacgaagcagctcagccctgcacgcgaatcggaacgtctgttgcagct 3600  
Db 3541 ttgtacttgacgaagcagctcagccctgcacgcgaatcggaacgtctgttgcagct 3600  
QY 3601 ggcctcgatgagccatccacagccgaacgaataagcagtgccgacccagcttccact 3660  
Db 3601 ggcctcgatgagccatccacagccgaacgaataagcagtgccgacccagcttccact 3660  
QY 3661 attcggaaatgttgatgttatttttggtttggccaaagggagaaatcgccgaacgggcaact 3720  
Db 3661 attcggaaatgttgatgttatttttggtttggccaaagggagaaatcgccgaacgggcaact 3720  
QY 3721 cagcgcgaactacaacgactgagagagaaatattacagagatgttttggcacatcttta 3780  
Db 3721 cagcgcgaactacaacgactgagagagaaatattacagagatgttttggcacatcttta 3780  
QY 3781 gaccaagcatga 3792  
Db 3781 gaccaagcatga 3792

RESULT 5  
US-09-677-488A-10  
; Sequence 10, Application US/09677488A  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 35718/204100  
; CURRENT APPLICATION NUMBER: US/09/677,488A  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 1999-07-12

; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 3792  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)---(0)  
; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA  
US-09-677-488A-10

Query Match 100.0%; Score 3791.2; DB 26; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgacagatgaatcgagaaacctgacaaaccagaatggcagtgagtcgtctctcacac 60  
DB 1 atgacagatgaatcgagaaacctgacaaaccagaatggcagtgagtcgtctctcacac 60  
QY 61 cctccccagaagaagaaaccgaagcagatattcagactatctcagaaattcagatat 120  
DB 61 cctccccagaagaagaaaccgaagcagatattcagactatctcagaaattcagatat 120  
QY 121 gccgacaaatcacactggactcctcaatgtcatcgctgcgtcactctgcgcacatcgatccggg 180  
DB 121 gccgacaaatcacactggactcctcaatgtcatcgctgcgtcactctgcgcacatcgatccggg 180  
QY 181 gctccctctctgtatcgatcactcttcgactagcttcacacaaagaattcaacaattac 240  
DB 181 gctccctctctgtatcgatcactcttcgactagcttcacacaaagaattcaacaattac 240  
QY 241 aattcggggcagggagtcgaagcgttcaaggccgatgagatcttcgtctctgtgg 300  
DB 241 aattcggggcagggagtcgaagcgttcaaggccgatgagatcttcgtctctgtgg 300  
QY 301 ttcgtctacactctttatgggaattgttcctcaogtcagctttccacggctgccattacc 360  
DB 301 ttcgtctacactctttatgggaattgttcctcaogtcagctttccacggctgccattacc 360  
QY 361 attcagctatagaaacctgaactcttcgacgagtggttctcttgatgacactctcgg 420  
DB 361 attcagctatagaaacctgaactcttcgacgagtggttctcttgatgacactctcgg 420  
QY 421 caagaggtctggcattctgcagaagcagcaatggagcaaatgccactcargtcactacc 480  
DB 421 caagaggtctggcattctgcagaagcagcaatggagcaaatgccactcargtcactacc 480  
QY 481 aatggcaacgtatatacaaacaggtattgccgagaaattggctttaccgtgcaggcactt 540  
DB 481 aatggcaacgtatatacaaacaggtattgccgagaaattggctttaccgtgcaggcactt 540  
QY 541 tcaatgtctctctcatttggctcgttttggcgttcacgtcagtgaaagctgatttaac 600  
DB 541 tcaatgtctctctcatttggctcgttttggcgttcacgtcagtgaaagctgatttaac 600  
QY 601 accatgtccgtcatccctgcatttcttcctgggtcacccgcaatctgcatagaattgatgcc 660  
DB 601 accatgtccgtcatccctgcatttcttcctgggtcacccgcaatctgcatagaattgatgcc 660  
QY 661 gctcagagggccaggaatcacaggatctactcacgcgcgtctcctcgcagagaagtc 720  
DB 661 gctcagagggccaggaatcacaggatctactcacgcgcgtctcctcgcagagaagtc 720  
QY 721 ttatcatcatcagcagatcctatgctttctacgcccacagaagaataatggtcgaaaaatat 780  
DB 721 ttatcatcatcagcagatcctatgctttctacgcccacagaagaataatggtcgaaaaatat 780  
QY 781 gatgtctttttgcagcaagcacaccacgaaggaaggaagaataatcgcgaataatggsgtc 840  
DB 781 gatgtctttttgcagcaagcacaccacgaaggaaggaagaataatcgcgaataatggsgtc 840

QY 841 ttgtctctcaactgagtgacttttgcattttagctgctgctatcgactgacctttttgaaagct 900  
DB 841 ttgtctctcaactgagtgacttttgcattttagctgctgctatcgactgacctttttgaaagct 900  
QY 901 tttegcatgtatcagaatggcgaggttggcgacgttggcgaagcttcttactgtgtccct 960  
DB 901 tttegcatgtatcagaatggcgaggttggcgacgttggcgaagcttcttactgtgtccct 960  
QY 961 tccgtcaccttcagcagcagctccatctcaatgttgcgtcttcaggttccagtcgctttac 1020  
DB 961 tccgtcaccttcagcagcagctccatctcaatgttgcgtcttcaggttccagtcgctttac 1020  
QY 1021 caacgcccatcttcgctccgcaattattcagtatcatctgacaaaccacgcagctcgac 1080  
DB 1021 caacgcccatcttcgctccgcaattattcagtatcatctgacaaaccacgcagctcgac 1080  
QY 1081 cctctcgaccttcttgaaagcagcagagggctgcctaggtccaatttgagatcccaaac 1140  
DB 1081 cctctcgaccttcttgaaagcagcagagggctgcctaggtccaatttgagatcccaaac 1140  
QY 1141 ctggcatttgcctaccctccgaccatctgccaaagcttctcgagatcttcaacttgaca 1200  
DB 1141 ctggcatttgcctaccctccgaccatctgccaaagcttctcgagatcttcaacttgaca 1200  
QY 1201 attcagctggcagaagcagggccctcgtcgttgatctcaggttagcgcgaacacaaatg 1260  
DB 1201 attcagctggcagaagcagggccctcgtcgttgatctcaggttagcgcgaacacaaatg 1260  
QY 1261 gtcggtctcaactgaacgggtggtatctgccagtttcggggaggatattacttgatgggttg 1320  
DB 1261 gtcggtctcaactgaacgggtggtatctgccagtttcggggaggatattacttgatgggttg 1320  
QY 1321 gaactgggacatacaaatgtgaaatggctggaagcgcgcatctgcctcgttcaacagga 1380  
DB 1321 gaactgggacatacaaatgtgaaatggctggaagcgcgcatctgcctcgttcaacagga 1380  
QY 1381 cctgtgtgttctcgtggcacaattctccagaacaattgcgaacacgggttctatggatgagca 1440  
DB 1381 cctgtgtgttctcgtggcacaattctccagaacaattgcgaacacgggttctatggatgagca 1440  
QY 1441 cgagatctgcctcgcgaacaaacaaatggagcttgcgaacaaagcttgcgaacacgacat 1500  
DB 1441 cgagatctgcctcgcgaacaaacaaatggagcttgcgaacaaagcttgcgaacacgacat 1500  
QY 1501 ggcgacgtgttcatttaagtgcgttcgaagcgttgcgaagcgttgcgaagcgttcgaagcgc 1560  
DB 1501 ggcgacgtgttcatttaagtgcgttcgaagcgttgcgaagcgttgcgaagcgttcgaagcgc 1560  
QY 1561 ggagccttgagtgagggtcaacgcacaacgaattgcgaatcgcaagagtatcatatcgat 1620  
DB 1561 ggagccttgagtgagggtcaacgcacaacgaattgcgaatcgcaagagtatcatatcgat 1620  
QY 1621 cccaagatcctcgttactcgtatgaagctaccagcgccttgaccgcagggcaggaagtg 1680  
DB 1621 cccaagatcctcgttactcgtatgaagctaccagcgccttgaccgcagggcaggaagtg 1680  
QY 1681 gtcagagggtcgttcgaacccagtgctcaaaagacgcgaactcttgcctatgcgcacaa 1740  
DB 1681 gtcagagggtcgttcgaacccagtgctcaaaagacgcgaactcttgcctatgcgcacaa 1740  
QY 1741 ctacgcactctcaaaaagtgcctggcaacatcgagctcatttcccagggggaataatcgtcgag 1800  
DB 1741 ctacgcactctcaaaaagtgcctggcaacatcgagctcatttcccagggggaataatcgtcgag 1800  
QY 1801 caaggcacacaccacgaattgatctcggctgtcattaccgcccacttggtggtcgca 1860  
DB 1801 caaggcacacaccacgaattgatctcggctgtcattaccgcccacttggtggtcgca 1860  
QY 1861 caggacctcggggtcgacgaacaaacagacatgagagacctctgcacgaagaaggcagca 1920  
DB 1861 caggacctcggggtcgacgaacaaacagacatgagagacctctgcacgaagaaggcagca 1920  
QY 1921 cgagagactgctggtgacacgaccggcacttgagcgcgactcacaccactgcatcatctcaa 1980



; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA  
US-09-677-682-10

Query Match 100.0%; Score 3791.2; DB 26; Length 3792;  
Best Local Similarity 100.0%; pred. NO. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OTHER INFORMATION: p-glycoprotein, fully spliced cDNA									
US-09-677-682-10									
Query Match 100.0%; Score 3791.2; DB 26; Length 3792;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	atgagcagtaacatcgagaaacccctgcacaaacaaagatggcagtgagtcctcacac	60						
Db	1	atgagcagtaacatcgagaaacccctgcacaaacaaagatggcagtgagtcctcacac	60						
Qy	61	ctctcccccgaagaaacgaagcgagttttcagactatctcagaatcttcagatat	120						
Db	61	ctctcccccgaagaaacgaagcgagttttcagactatctcagaatcttcagatat	120						
Qy	121	gccgacaaatcagactcgactcgaatcgaatcgctcgcctcgcgcacatcgagtcgg	180						
Db	121	gccgacaaatcagactcgactcgaatcgaatcgctcgcctcgcgcacatcgagtcgg	180						
Qy	181	gtctccctctctgactgcatctcgtatctcgttagcttcacaaacgaagttcaacaattac	240						
Db	181	gtctccctctctgactgcatctcgtatctcgttagcttcacaaacgaagttcaacaattac	240						
Qy	241	aattcgccgcaggaagctcgaagcgttcgaagccgaatggatcatttgcctgtgag	300						
Db	241	aattcgccgcaggaagctcgaagcgttcgaagccgaatggatcatttgcctgtgag	300						
Qy	301	ttcgtctaacctttatttgggaagtttgtctctcagtaagtttccacggctgcattacc	360						
Db	301	ttcgtctaacctttatttgggaagtttgtctctcagtaagtttccacggctgcattacc	360						
Qy	361	atttcagctatcacgaacactcgaactcttcagcagtggttccttgaatgcaccttgagg	420						
Db	361	atttcagctatcacgaacactcgaactcttcagcagtggttccttgaatgcaccttgagg	420						
Qy	421	caagaggtctggcattctcgacaacgagcgaatggagcaatcgccactcargtcaactacc	480						
Db	421	caagaggtctggcattctcgacaacgagcgaatggagcaatcgccactcargtcaactacc	480						
Qy	481	aatggcaacggtatacaaacagggtattgcccagagaaattggtcttaccgtgcaggcaactt	540						
Db	481	aatggcaacggtatacaaacagggtattgcccagagaaattggtcttaccgtgcaggcaactt	540						
Qy	541	tcaatgtctttctgcatttgtgtgcgtttggtcgtctcagtggaagctagatttcaatc	600						
Db	541	tcaatgtctttctgcatttgtgtgcgtttggtcgtctcagtggaagctagatttcaatc	600						
Qy	601	accatgtccgtatcccttgcaatttctcgtgtcaccggcacttgcatagcaattgatgcc	660						
Db	601	accatgtccgtatcccttgcaatttctcgtgtcaccggcacttgcatagcaattgatgcc	660						
Qy	661	gtccaggaggccagatcacagagatctactcacgcgcgcgtactctcgaaagaagtc	720						
Db	661	gtccaggaggccagatcacagagatctactcacgcgcgcgtactctcgaaagaagtc	720						
Qy	721	ttatcatctccgacagatccatgttcttaagccacagagaataatggtcgaaaaatat	780						
Db	721	ttatcatctccgacagatccatgttcttaagccacagagaataatggtcgaaaaatat	780						
Qy	781	gatgtctttttgcagacagcacacagaaggggaagaagaaatcgccaaataatggsgtc	840						
Db	781	gatgtctttttgcagacagcacacagaaggggaagaagaaatcgccaaataatggsgtc	840						
Qy	841	ttggtctaacctgagtaacttttgcatctacgtctgctacgtcactggcccttttggaaggt	900						
Db	841	ttggtctaacctgagtaacttttgcatctacgtctgctacgtcactggcccttttggaaggt	900						
Qy	901	tttcgcagtatcacaaatggcgaggttgcgcaggttggcgaagctcttactgttgcctt	960						
Db	901	tttcgcagtatcacaaatggcgaggttgcgcaggttggcgaagctcttactgttgcctt	960						
Qy	961	tcctgcacttagcagcaagctccatctcaatgcttgcgcttcaaggttcaagtcgtttac	1020						
Db	961	tcctgcacttagcagcaagctccatctcaatgcttgcgcttcaaggttcaagtcgtttac	1020						
Qy	1021	caacgcgcgcatcttcgggtctccgaattatttcagttatcattgacaaacccacgcagctcgac	1080						
Db	1021	caacgcgcgcatcttcgggtctccgaattatttcagttatcattgacaaacccacgcagctcgac	1080						
Qy	1081	ctctcgcagccctcttggaaagcagcagaggtggtcctcagatcctaaattgagatcccaaac	1140						
Db	1081	ctctcgcagccctcttggaaagcagcagaggtggtcctcagatcctaaattgagatcccaaac	1140						
Qy	1141	ctggcattttgcttaccctcccgacacattctgcccaagtaacttcagagatttcaacttgaca	1200						
Db	1141	ctggcattttgcttaccctcccgacacattctgcccaagtaacttcagagatttcaacttgaca	1200						
Qy	1201	attccagcttggcagaagcagcgcctctcgtgtgcatcaggtagcgcaaaagacacaaatg	1260						
Db	1201	attccagcttggcagaagcagcgcctctcgtgtgcatcaggtagcgcaaaagacacaaatg	1260						
Qy	1261	gtcgggttacttgaacgattggtatctcccaagtttcggagaggaattacttgcattggtgtg	1320						
Db	1261	gtcgggttacttgaacgattggtatctcccaagtttcggagaggaattacttgcattggtgtg	1320						
Qy	1321	gaactgggacaatacaatgtgaatggctggaagacgcgcattcgccctcgtttcaacagggaa	1380						
Db	1321	gaactgggacaatacaatgtgaatggctggaagacgcgcattcgccctcgtttcaacagggaa	1380						
Qy	1381	ctgtgtgtttgttggcacaatatcttcagaaatctgccaaacgggtttcaatggatgagcaaa	1440						
Db	1381	ctgtgtgtttgttggcacaatatcttcagaaatctgccaaacgggtttcaatggatgagcaaa	1440						
Qy	1441	cgagaactgcctcgcgcaaaaaacaaatggagcttctgcnaaaagcttgcnaaagccagcaaat	1500						
Db	1441	cgagaactgcctcgcgcaaaaaacaaatggagcttctgcnaaaagcttgcnaaagccagcaaat	1500						
Qy	1501	ggcgacgttctcaatcaatgagcttccgaacggttatgagactgaagtgtgcgcagcagcgc	1560						
Db	1501	ggcgacgttctcaatcaatgagcttccgaacggttatgagactgaagtgtgcgcagcagcgc	1560						
Qy	1561	ggagccttgagtgagggtcaacgcacaaagaaattgcaatcgacagaagtaacatctcggaat	1620						
Db	1561	ggagccttgagtgagggtcaacgcacaaagaaattgcaatcgacagaagtaacatctcggaat	1620						
Qy	1621	cccaagatcctgttactcgatgaagctaccagcgcctctgaccgcgaaggcggaagaagtg	1680						
Db	1621	cccaagatcctgttactcgatgaagctaccagcgcctctgaccgcgaaggcggaagaagtg	1680						
Qy	1681	gtccaggaggccttgaaccgaggtgtccaaagacgcgcactcttttgggtcattgcccacaaa	1740						
Db	1681	gtccaggaggccttgaaccgaggtgtccaaagacgcgcactcttttgggtcattgcccacaaa	1740						
Qy	1741	ctagccacttgcacaaagtgtcggcaacatcgactcatttcccaaggaggaataatcgctcga	1800						
Db	1741	ctagccacttgcacaaagtgtcggcaacatcgactcatttcccaaggaggaataatcgctcga	1800						
Qy	1801	caaggcacacacacagaattgatcgaaattcggtctgctcattacgcgcgcacatcggtcgca	1860						
Db	1801	caaggcacacacacagaattgatcgaaattcggtctgctcattacgcgcgcacatcggtcgca	1860						
Qy	1861	caggacactcggggtgcagcaaacacagaacatgagaagacccctgcgcgaagaaggcagca	1920						

2041	Db	tycatcctaatactgtttactagacacadaaaaaattctctactggtgctctctgtgtgtcaca	2100
2101	QY	ataacggtctctgatatctgcggcgccacatttccagagcaagccctttgtttttcagagattg	2160
2101	Db	ataacggtctctgatatctgcggcgccacatttccagagcaagccctttgtttttcagagattg	2160
2161	QY	ctcactctcttcgagttgagttggtctacgcgcacagaaacggcagacattttatattctg	2220
2161	Db	ctcactctcttcgagttgagttggtctacgcgcacagaaacggcagacattttatattctg	2220
2221	QY	atgtctcttctgctgactctagaaaaactagtaggataattcacagattgctcgacaatcg	2280
2221	Db	atgtctcttctgctgactctagaaaaactagtaggataattcacagattgctcgacaatcg	2280
2281	QY	aacgttatttcaaaattgttcacccaatcgctatatcaagccgcaaatgttccaacagattctg	2340
2281	Db	aacgttatttcaaaattgttcacccaatcgctatatcaagccgcaaatgttccaacagattctg	2340
2341	QY	gatcaagacatcgaaactctcgacatacccgagcgaataattctggtctctcacatcgcaa	2400
2341	Db	gatcaagacatcgaaactctcgacatacccgagcgaataattctggtctctcacatcgcaa	2400
2401	QY	ctgtcagctctaccacgcgaattgcgaagttgatatacgaaataatttctcatttatctc	2460
2401	Db	ctgtcagctctaccacgcgaattgcgaagttgatatacgaaataatttctcatttatctc	2460
2461	QY	gtttctcggtcacaatcgctctgcagcagctctctaccactagcctatgagtgaacactggg	2520
2461	Db	gtttctcggtcacaatcgctctgcagcagctctctaccactagcctatgagtgaacactggg	2520
2521	QY	octggttggttctgtttggttgcaacttccaccctgcttttggctggtcactccaagaattcg	2580
2521	Db	octggttggttctgtttggttgcaacttccaccctgcttttggctggtcactccaagaattcg	2580
2581	QY	tctagagacaaactagaaagccggaactcggcaaaactttgcagaaagtctcgtggtctgc	2640
2581	Db	tctagagacaaactagaaagccggaactcggcaaaactttgcagaaagtctcgtggtctgc	2640
2641	QY	aagcgaagcagttacccgcgatccgcgcgatccgcgtctcatcttgcactctcgaagggccatgtct	2700
2641	Db	aagcgaagcagttacccgcgatccgcgcgatccgcgtctcatcttgcactctcgaagggccatgtct	2700
2701	QY	ccaacagtactcggacatgttgacaaagctcttgctaaagatcatccaagcttttgatt	2760
2701	Db	ccaacagtactcggacatgttgacaaagctcttgctaaagatcatccaagcttttgatt	2760
2761	QY	ggacgattgtttgttttccattgttgcacatctcagatggagtgtttggctattgcctggga	2820
2761	Db	ggacgattgtttgttttccattgttgcacatctcagatggagtgtttggctattgcctggga	2820
2821	QY	ttttgtatttcagtcgaataatttgcttcaggtcaggtacgacacaaactcaatttatatca	2880
2821	Db	ttttgtatttcagtcgaataatttgcttcaggtcaggtacgacacaaactcaatttatatca	2880
2881	QY	tcttcgtggcgctttgtttgctcggctccaaacgacgaccagctatttggcttaactccacg	2940
2881	Db	tcttcgtggcgctttgtttgctcggctccaaacgacgaccagctatttggcttaactccacg	2940
2941	QY	agttttaccaaagctcgttcgctgcggaactatactcttggctgcgacattgaagccg	3000
2941	Db	agttttaccaaagctcgttcgctgcggaactatactcttggctgcgacattgaagccg	3000
3001	QY	accttcgcgaaacggagagagacaaagaaaagccagatgggtggatgccctgtcgac	3060
3001	Db	accttcgcgaaacggagagagacaaagaaaagccagatgggtggatgccctgtcgac	3060
3061	QY	ctcagagacattgaaatccaggtatctcgtcaacggtgatttcggctcgagttctcgcggggtt	3120
3061	Db	ctcagagacattgaaatccaggtatctcgtcaacggtgatttcggctcgagttctcgcggggtt	3120
3121	QY	tcctagacaattcgcagcagacaaatttgtacttatgtggcgcttctggctgtggcaag	3180
3121	Db	tcctagacaattcgcagcagacaaatttgtacttatgtggcgcttctggctgtggcaag	3180

Qy	3181	tcaacgttgatcgcttgttcgaaagcatctctacgaccgcgaactcgggcgaatttcattt	3240
Db	3181	tcaacgttgatcgcttgttcgaaagcatctctacgaccgcgaactcgggcgaatttcattt	3240
Qy	3241	gcacacgagaatattgcagaaaatgcgcgcgcttgcacccggccatatgtctttggctc	3300
Db	3241	gcacacgagaatattgcagaaaatgcgcgcgcttgcacccggccatatgtctttggctc	3300
Qy	3301	caacaggaaccccaactttaccaaagctccgttcgcgagaatgtgacgttgccctcgaa	3360
Db	3301	caacaggaaccccaactttaccaaagctccgttcgcgagaatgtgacgttgccctcgaa	3360
Qy	3361	gccgaattatcagaagagctttgtcaagagacgcttccgcgaaggccaatctcttgattt	3420
Db	3361	gccgaattatcagaagagctttgtcgaagagacgcttccgcgaaggccaatctcttgattt	3420
Qy	3421	ttgtcatctctttaccagaaggctttgaacgcttgcggtccaacgaggatgcagtttc	3480
Db	3421	ttgtcatctctttaccagaaggctttgaacgcttgcggtccaacgaggatgcagtttc	3480
Qy	3481	tcggcgggcaacgacagcgatcgccatcgcaagagcatgtatcgcaatcccaagctg	3540
Db	3481	tcggcgggcaacgacagcgatcgccatcgcaagagcatgtatcgcaatcccaagctg	3540
Qy	3541	ttgtctacttgacgaagcgcgtcagccctcgcacgcgaatcggaaactcttggtccaagct	3600
Db	3541	ttgtctacttgacgaagcgcgtcagccctcgcacgcgaatcggaaactcttggtccaagct	3600
Qy	3601	gccctcgatgaggcatccacgagcccaacgacaatagcagttggcgaccgcactttccact	3660
Db	3601	gccctcgatgaggcatccacgagcccaacgacaatagcagttggcgaccgcactttccact	3660
Qy	3661	attcggaattatgatgttattttgtttgcgaacggaggaatcgcgcaaacggggcaact	3720
Db	3661	attcggaatgtcgatgtattttgtttgcgaacggaggaatcgcgcaaacggggcaact	3720
Qy	3721	caacgcgaactacaacgactcgaagagaatattacgagatgtgtttgggcaaatctttaa	3780
Db	3721	caacgcgaactacaacgactcgaagagaatattacgagatgtgtttgggcaaatctttaa	3780
Qy	3781	gacccaagcatga	3792
Db	3781	gacccaagcatga	3792

```

RESULT# 7
US-09-677-682A-10
; Sequence 10, Application US/09677682A
; GENERAL INFORMATION:
; APPLICANT: Duviack, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Grasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; TITLE OF INVENTION: Detoxification
; FILE REFERENCE: 35718/204101
; CURRENT APPLICATION NUMBER: US/09/677.682A
; CURRENT FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 3792
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA
US-09-677-682A-10

```



Query Match		100.0%	Score 3791.2	DB 26	Length 3792
Best Local Similarity		100.0%	Pred. No. 0		
Matches 3792		Conservative	0	Mismatches	0
		Indels	0	Gaps	0
Qy	1	atggcagatgaatcggagaaacctcgaccacaaacaaagatggcagtgagtcgtctcacac	60		
Db	1	atggcagatgaatcggagaaacctcgaccacaaacaaagatggcagtgagtcgtctcacac	60		
Qy	61	ccctcccccagaaagaaacaaagcagatatttcagactatctacgaattctcaagatat	120		
Db	61	ccctcccccagaaagaaacaaagcagatatttcagactatctacgaattctcaagatat	120		
Qy	121	gccgacaaatacagctgagctctcaatgtcatcgctcatctcgccatcgatccggg	180		
Db	121	gccgacaaatacagctgagctctcaatgtcatcgctcatctcgccatcgatccggg	180		
Qy	181	gtctccctctctgtatgtcgatcatcttcggtagcttcacaaagttcaacaattac	240		
Db	181	gtctccctctctgtatgtcgatcatcttcggtagcttcacaaagttcaacaattac	240		
Qy	241	aattcgggagcaggtcctgaagcgttaagccgagtggtgatacttcgttcgtg	300		
Db	241	aattcgggagcaggtcctgaagcgttaagccgagtggtgatacttcgttcgtg	300		
Qy	301	ttcgtctactctttattgggaagttgtcctcactgactggtttccacggctgcattac	360		
Db	301	ttcgtctactctttattgggaagttgtcctcactgactggtttccacggctgcattac	360		
Qy	361	atttcagctatcgaaccactcgactcttcgacgagtggttccttgatgcacttcggg	420		
Db	361	atttcagctatcgaaccactcgactcttcgacgagtggttccttgatgcacttcggg	420		
Qy	421	caagaggttcgcttcgacaaagcagagcaatgagcaatccacactcargtcaatcc	480		
Db	421	caagaggttcgcttcgacaaagcagagcaatgagcaatccacactcargtcaatcc	480		
Qy	481	aatggcaacgctatacaaacaggtattcccgagaaattggtcttaccgtgcaggcactt	540		
Db	481	aatggcaacgctatacaaacaggtattcccgagaaattggtcttaccgtgcaggcactt	540		
Qy	541	tcaatgtctttctgcatcttggtgctgttcgctcagtggaagctagctttaac	600		
Db	541	tcaatgtctttctgcatcttggtgctgttcgctcagtggaagctagctttaac	600		
Qy	601	accatgtccgtcaatccctgcaatttctgctacccgcatctcagatagcaattgagtc	660		
Db	601	accatgtccgtcaatccctgcaatttctgctacccgcatctcagatagcaattgagtc	660		
Qy	661	gctcaggagccaggaataccaggaatctactcagcgcgcgtgtcctcgcagaagaagtc	720		
Db	661	gctcaggagccaggaataccaggaatctactcagcgcgcgtgtcctcgcagaagaagtc	720		
Qy	721	ttatcatcatcccgagacagtcctgttctacgcccagagaaaaatggctcgaaaaatat	780		
Db	721	ttatcatcatcccgagacagtcctgttctacgcccagagaaaaatggctcgaaaaatat	780		
Qy	781	gatgtcttttgcagcaagcaccagagagggagagaaatcgccaaataatgsgtc	840		
Db	781	gatgtcttttgcagcaagcaccagagagggagagaaatcgccaaataatgsgtc	840		
Qy	841	ttgtttcactgagtaacttttgaatttgaactgtctatgcactgcgcttttggaaaggt	900		
Db	841	ttgtttcactgagtaacttttgaatttgaactgtctatgcactgcgcttttggaaaggt	900		
Qy	901	tttcgcatgtatcagaatggagaggttgcgcagcttggcaagcttttactgtgtccctt	960		
Db	901	tttcgcatgtatcagaatggagaggttgcgcagcttggcaagcttttactgtgtccctt	960		
Qy	961	tcgctcactttagcagcagctccatctcaagcttgcgcttcaggttcagtcgtttac	1020		
Db	961	tcgctcactttagcagcagctccatctcaagcttgcgcttcaggttcagtcgtttac	1020		

Qy	1021	caacgcgcacatcttcggtctccgaattattcagtatcattgacaaacccacgcagctcgac	1080
Db	1021	caacgcgcacatcttcggtctccgaattattcagtatcattgacaaacccacgcagctcgac	1080
Qy	1081	ccctcgcaccttcttggaagcagccagagggctcctcaggtcctcaggttcaggttcgagatccaaac	1140
Db	1081	ccctcgcaccttcttggaagcagccagagggctcctcaggtcctcaggttcaggttcgagatccaaac	1140
Qy	1141	ctggcatttgcctacccctcccgacctctgcgacagttacttgcgatttcgagattgaca	1200
Db	1141	ctggcatttgcctacccctcccgacctctgcgacagttacttgcgatttcgagattgaca	1200
Qy	1201	attccagctggcagacgagccctcgtcgtgtgcatcaggtagcggcaaaagcacaaatg	1260
Db	1201	attccagctggcagacgagccctcgtcgtgtgcatcaggtagcggcaaaagcacaaatg	1260
Qy	1261	gtcggcttacttgaaacgggtggtatctgcgccagttcgggagagatatctacttgatgggtt	1320
Db	1261	gtcggcttacttgaaacgggtggtatctgcgccagttcgggagagatatctacttgatgggtt	1320
Qy	1321	gaactgggacaaatacaatgtaaatggtgagaaagccgcatctgcctcttcaaacagaa	1380
Db	1321	gaactgggacaaatacaatgtaaatggtgagaaagccgcatctgcctcttcaaacagaa	1380
Qy	1381	cctgtgtgttctcgtggcaaatcttccgaaacatttgcacacgggttctcatgagacaa	1440
Db	1381	cctgtgtgttctcgtggcaaatcttccgaaacatttgcacacgggttctcatgagacaa	1440
Qy	1441	cgagatctgctcgcgaaacaaatggagcttgcgaaagcttgcgaaagccagcgaat	1500
Db	1441	cgagatctgctcgcgaaacaaatggagcttgcgaaagcttgcgaaagccagcgaat	1500
Qy	1501	ggcagcgtgttcattatgagcttcgcaacggttatgagactgaagttgagcagagcc	1560
Db	1501	ggcagcgtgttcattatgagcttcgcaacggttatgagactgaagttgagcagagcc	1560
Qy	1561	ggcgaccttgagtgaggttcaacgcaacgaattgcaatcgcaacgaagtatcatctcgat	1620
Db	1561	ggcgaccttgagtgaggttcaacgcaacgaattgcaatcgcaacgaagtatcatctcgat	1620
Qy	1621	cccaagatcctgttactcgtatgaagctacagcgccttgcgcaagggcggaagagtg	1680
Db	1621	cccaagatcctgttactcgtatgaagctacagcgccttgcgcaagggcggaagagtg	1680
Qy	1681	gtccagagccttgaaacgaatgtccaaagaccgcatcttggctcattggccacaa	1740
Db	1681	gtccagagccttgaaacgaatgtccaaagaccgcatcttggctcattggccacaa	1740
Qy	1741	ctaggccactgtcaaaagtctggcaacatcgagtcatttcccagggggaaatcgtcgag	1800
Db	1741	ctaggccactgtcaaaagtctggcaacatcgagtcatttcccagggggaaatcgtcgag	1800
Qy	1801	caaggcacacacacgaattgatcgaattcggctgtcattacgcgcgactggtgctgca	1860
Db	1801	caaggcacacacacgaattgatcgaattcggctgtcattacgcgcgactggtgctgca	1860
Qy	1861	caggacctcggggctgacgaacacaaagaaatgagagacccctgcagaaagagcaga	1920
Db	1861	caggacctcggggctgacgaacacaaagaaatgagagacccctgcagaaagagcaga	1920
Qy	1921	cgagagctgtcgttgatgaacgacccgcatcttgagcgcactcaacacactgcacactca	1980
Db	1921	cgagagctgtcgttgatgaacgacccgcatcttgagcgcactcaacacactgcacactca	1980
Qy	1981	gctgggagactggagagcggaaggtgcggttcgggactttgggtcactcgtcctctaaa	2040
Db	1981	gctgggagactggagagcggaaggtgcggttcgggactttgggtcactcgtcctctaaa	2040
Qy	2041	tgcactcctaactcatgtctcagcaacaaataatctctactggtgctctctgtgtgcaaca	2100
Db	2041	tgcactcctaactcatgtctcagcaacaaataatctctactggtgctctctgtgtgcaaca	2100







Db 2101 ataaagcttctgatatgcgagccacatttccagacaaagccctttgtttctcgagattg 2160  
QY 2161 ctcaactgtcttcgaattgaattgctatgcgcacaggaacgagcaacttttatattctg 2220  
Db 2161 ctcaactgtcttcgaattgaattgctatgcgcacaggaacgagcaacttttatattctg 2220  
QY 2221 atgtctcttctgttggctctagaaatcttagtagatatttcaagatttgcgtgacatgc 2280  
Db 2221 atgtctcttctgttggctctagaaatcttagtagatatttcaagatttgcgtgacatgc 2280  
QY 2281 aacgttatttcaacagttgtcaaccatcgctatcaagccgcgaatttccaacagtagtgc 2340  
Db 2281 aacgttatttcaacagttgtcaaccatcgctatcaagccgcgaatttccaacagtagtgc 2340  
QY 2341 gatacagacatcgaaactctcgacatcccgagcaaatcttggctctcacaatcgcaa 2400  
Db 2341 gatacagacatcgaaactctcgacatcccgagcaaatcttggctctcacaatcgcaa 2400  
QY 2401 ctgtcagctctaccacagcagttgcgaagattgatatcgacaaattttctcatttatc 2460  
Db 2401 ctgtcagctctaccacagcagttgcgaagattgatatcgacaaattttctcatttatc 2460  
QY 2461 gttgtcgttcaacatcgctctcgagcagtgctctaccactagcctatggtgaaactggg 2520  
Db 2461 gttgtcgttcaacatcgctctcgagcagtgctctaccactagcctatggtgaaactggg 2520  
QY 2521 cctggtgttattgtttgttgatctccaccctgcttttggctgagctacatcagaattcg 2580  
Db 2521 cctggtgttattgtttgttgatctccaccctgcttttggctgagctacatcagaattcg 2580  
QY 2581 tctagacagcagctagaagccggaactcggcaaaacttgcgaagaagtgctggcttcgc 2640  
Db 2581 tctagacagcagctagaagccggaactcggcaaaacttgcgaagaagtgctggcttcgc 2640  
QY 2641 aagcgaagcagttacccgcagctccggaccgtctctatcttggactctcgaagccatgttct 2700  
Db 2641 aagcgaagcagttacccgcagctccggaccgtctctatcttggactctcgaagccatgttct 2700  
QY 2701 ccaacagctactcgacatcttgagcaaggtcttgaatcattcgaagcttttggttt 2760  
Db 2701 ccaacagctactcgacatcttgagcaaggtcttgaatcattcgaagcttttggttt 2760  
QY 2761 ggaacatgttttggatttcaactgtctcaactcgatagagatttttggctatgcctcggga 2820  
Db 2761 ggaacatgttttggatttcaactgtctcaactcgatagagatttttggctatgcctcggga 2820  
QY 2821 ttttggattgcagtgataattggcttcaggtgagtagcagacacaaactcaattttatca 2880  
Db 2821 ttttggattgcagtgataattggcttcaggtgagtagcagacacaaactcaattttatca 2880  
QY 2881 tcttcgtggcgttttggcttcaggttcgaagcagccagctatttggcttactccag 2940  
Db 2881 tcttcgtggcgttttggcttcaggttcgaagcagccagctatttggcttactccag 2940  
QY 2941 agtttaccagcctcgttcgctgcgaactatactctgcgtcgcgaacattgaaacgcg 3000  
Db 2941 agtttaccagcctcgttcgctgcgaactatactctgcgtcgcgaacattgaaacgcg 3000  
QY 3001 accatccgaaacgaggaagaaacaaagaaagcccggtgggtggatgcctcgtgcac 3060  
Db 3001 accatccgaaacgaggaagaaacaaagaaagcccggtgggtggatgcctcgtgcac 3060  
QY 3061 ctcgaggaattgaattcagttatcgtcaacgtgattcgtcgtcgtcgtcgtcgtcgtt 3120  
Db 3061 ctcgaggaattgaattcagttatcgtcaacgtgattcgtcgtcgtcgtcgtcgtt 3120  
QY 3121 tcaatgaacatcgacaggaacaaatttgaattatgtgggcgttcttggctgtgggaag 3180  
Db 3121 tcaatgaacatcgacaggaacaaatttgaattatgtgggcgttcttggctgtgggaag 3180  
QY 3181 tcaacgttgcgttttgcgaacagatttctacgacccacacctcgggcgaatttcattt 3240  
Db 3181 tcaacgttgcgttttgcgaacagatttctacgacccgacacctcgggcgaatttcattt 3240

QY 3241 gcacacgaaatattgcagaaatgtccgcgcttgcgcgcttaccgcggccatatgtcttggctc 3300  
Db 3241 gcacacgaaatattgcagaaatgtccgcgcttgcgcgcttaccgcggccatatgtcttggctc 3300  
QY 3301 caacaggaacccacactttaccagaggtccgttgcgagaaatgtgacgttggccctcgaa 3360  
Db 3301 caacaggaacccacactttaccagaggtccgttgcgagaaatgtgacgttggccctcgaa 3360  
QY 3361 gccgaattatcatcagaagagctttgtcaaggagccttcccgcgaagccaatgtcttggatt 3420  
Db 3361 gccgaattatcatcagaagagctttgtcaaggagccttcccgcgaagccaatgtcttggatt 3420  
QY 3421 ttgtcattctttaccagaaggtcttgaacgcttgcgctcgaacgagggatgcagttc 3480  
Db 3421 ttgtcattctttaccagaaggtcttgaacgcttgcgctcgaacgagggatgcagttc 3480  
QY 3481 tccggcgggcaacacagcagcggtccatccgaagagcatttgcgaatccaaagctg 3540  
Db 3481 tccggcgggcaacacagcagcggtccatccgaagagcatttgcgaatccaaagctg 3540  
QY 3541 ttgtcattcttgaacgagcagctcagcctcgcacacgcaatcggaacgtctgttcaagct 3600  
Db 3541 ttgtcattcttgaacgagcagctcagcctcgcacacgcaatcggaacgtctgttcaagct 3600  
QY 3601 gcctcgtatgagcgtccacacgagcgaacgaataagtagtgcgcacgactttccact 3660  
Db 3601 gcctcgtatgagcgtccacacgagcgaacgaataagtagtgcgcacgactttccact 3660  
QY 3661 attcgaattgtgatttatttttgcgttgcgaacgaggaatccgaaacgagggcact 3720  
Db 3661 attcgaattgtgatttatttttgcgttgcgaacgaggaatccgaaacgagggcact 3720  
QY 3721 cagcggaaactacaacgactgaggaagatattacagagatgttggcacaattttta 3780  
Db 3721 cagcggaaactacaacgactgaggaagatattacagagatgttggcacaattttta 3780  
QY 3781 gaccaagcatga 3792  
Db 3781 gaccaagcatga 3792

RESULT 9  
US-09-882-694-10  
; Sequence 10, Application US/09882694  
; GENERAL INFORMATION:  
; APPLICANT: DUVICK, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; TITLE OF INVENTION: Detoxification  
; FILE REFERENCE: 35718/208255  
; CURRENT APPLICATION NUMBER: US/09/882,694  
; CURRENT FILING DATE: 2001-06-12  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 3792  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA  
US-09-882-694-10

Query Match: 100.0%; Score 3791.2; DB 33; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 1	atgacagatgaatcgagagaaactcgacacaaacgaagatggcagtgagtcgtctcac 60
Db 1	atgacagatgaatcgagagaaactcgacacaaacgaagatggcagtgagtcgtctcac 60
Qy 61	ctctcccccagaaagaaacccgaagcgagatattcagactactctacgaattctcagatat 120
Db 61	ctctcccccagaaagaaacccgaagcgagatattcagactactctacgaattctcagatat 120
Qy 121	gcgcacaaatcacactggactctcaatctcatcgctctcatctcgccatcggaatcg 180
Db 121	gcgcacaaatcacactggactctcaatctcatcgctctcatctcgccatcggaatcg 180
Qy 181	gttccctctctctgtagctgatctatcttcgttagcttccacaaagttccaaatctac 240
Db 181	gttccctctctctgtagctgatctatcttcgttagcttccacaaagttccaaatctac 240
Qy 241	aattcggggcagcgggagtcctgaagcgttccaaagccgagtggtgattctctctgtg 300
Db 241	aattcggggcagcgggagtcctgaagcgttccaaagccgagtggtgattctctctgtg 300
Qy 301	ttcgtctacactctttatgggaagttgtctcactgagtttccacggctgcattaac 360
Db 301	ttcgtctacactctttatgggaagttgtctcactgagtttccacggctgcattaac 360
Qy 361	atttcagctatcagaaacactcgaaactcttcgacgagtggttcttgaatgcaactgg 420
Db 361	atttcagctatcagaaacactcgaaactcttcgacgagtggttcttgaatgcaactgg 420
Qy 421	caagaggtcggcattctcagaaacagagcaatggagcaatcgccactcargtcaactacc 480
Db 421	caagaggtcggcattctcagaaacagagcaatggagcaatcgccactcargtcaactacc 480
Qy 481	aattggaacgtatcaaaacagttattcccgagaaattgctcttaccgagcaggaactt 540
Db 481	aattggaacgtatcaaaacagttattcccgagaaattgctcttaccgagcaggaactt 540
Qy 541	tcaatgtctttctcgtatcttgctgtcttggctcgtcgaagtgagtgagtttaac 600
Db 541	tcaatgtctttctcgtatcttgctgtcttggctcgtcgaagtgagtgagtttaac 600
Qy 601	accatgtccgtcactccctgcaattctcgtgctcaccggaatctcagcaatgtatgcc 660
Db 601	accatgtccgtcactccctgcaattctcgtgctcaccggaatctcagcaatgtatgcc 660
Qy 661	gctcaggaggcaggaatcaccaggtatctcagcgcgcgtgtcctcgagagaagtc 720
Db 661	gctcaggaggcaggaatcaccaggtatctcagcgcgcgtgtcctcgagagaagtc 720
Qy 721	ttatcattcattcagagatcattcttctcagcgcgcgtgtcctcgagagaagtc 780
Db 721	ttatcattcattcagagatcattcttctcagcgcgcgtgtcctcgagagaagtc 780
Qy 781	gatgtcttttgcagcaagcaccacaaaggaaggaagaaatgccaaataatgsgtc 840
Db 781	gatgtcttttgcagcaagcaccacaaaggaaggaagaaatgccaaataatgsgtc 840
Qy 841	ttgtctcactcagtagtacttttgcatttaccgtctatcgcactggccttttggaaagt 900
Db 841	ttgtctcactcagtagtacttttgcatttaccgtctatcgcactggccttttggaaagt 900
Qy 901	tttcgcatatcaagaatggcaggttgcgcagattggcgaagttcttactgttgcctt 960
Db 901	tttcgcatatcaagaatggcaggttgcgcagattggcgaagttcttactgttgcctt 960
Qy 961	ttcgtcaactcagagcagcagcttccatctcaatgttgcgctcaggttccgtttac 1020
Db 961	ttcgtcaactcagagcagcagcttccatctcaatgttgcgctcaggttccgtttac 1020
Qy 1021	caacgcgcagctctcgggtccgaaattatcagtatcattgacaaacccagcagctcgac 1080
Db 1021	caacgcgcagctctcgggtccgaaattatcagtatcattgacaaacccagcagctcgac 1080

Qy 1081	ctctcagacccctctctggaagcagccagagggctgctaggtccaaattgagatccaaac 1140
Db 1081	ctctcagacccctctctggaagcagccagagggctgctaggtccaaattgagatccaaac 1140
Qy 1141	ctgacatttgcctacccctccgacattctgccaagctacttcgagatttcaacttgaca 1200
Db 1141	ctgacatttgcctacccctccgacattctgccaagctacttcgagatttcaacttgaca 1200
Qy 1201	attccagctggcagacgcgcctcgtcggtgcatcaggtagcgcaaaagcacaatg 1260
Db 1201	attccagctggcagacgcgcctcgtcggtgcatcaggtagcgcaaaagcacaatg 1260
Qy 1261	gtcggcttacttgaaacgggtgtatctgcccagttcggggaggatattacttgatgggttg 1320
Db 1261	gtcggcttacttgaaacgggtgtatctgcccagttcggggaggatattacttgatgggttg 1320
Qy 1321	gaactgggacaaacacaaatgtgaaatggctgagaagccgcatctgctctgcttcaacagaa 1380
Db 1321	gaactgggacaaacacaaatgtgaaatggctgagaagccgcatctgctctgcttcaacagaa 1380
Qy 1381	ctctgtgtgttctcgtggcacaatcttccagaacattgccaacgggttccatgagcaaa 1440
Db 1381	ctctgtgtgttctcgtggcacaatcttccagaacattgccaacgggttccatgagcaaa 1440
Qy 1441	cgagatcgtcctcgcaaaacaaatggagctgtgcaaaagcttgcgaagccagcaat 1500
Db 1441	cgagatcgtcctcgcaaaacaaatggagctgtgcaaaagcttgcgaagccagcaat 1500
Qy 1501	ggcagcgttctcattatagcttccgaacggtctatgagactggaagtggcgagcgacc 1560
Db 1501	ggcagcgttctcattatagcttccgaacggtctatgagactggaagtggcgagcgacc 1560
Qy 1561	ggagccttgagtgaggttcaacgacaaatgtgcaatcgcaacgagtatcatatcgat 1620
Db 1561	ggagccttgagtgaggttcaacgacaaatgtgcaatcgcaacgagtatcatatcgat 1620
Qy 1621	cccaagatcctgttactcgtatgaagctacagcgccttgacccgagggcgagaaagt 1680
Db 1621	cccaagatcctgttactcgtatgaagctacagcgccttgacccgagggcgagaaagt 1680
Qy 1681	gtcagagagccttgaaacggagtgccaagacgcactactttggtcatttgcacacaa 1740
Db 1681	gtcagagagccttgaaacggagtgccaagacgcactactttggtcatttgcacacaa 1740
Qy 1741	ctagcacttcaaaagtgtcgcaacatcgagctcatttccagggggaagaaatcgctcag 1800
Db 1741	ctagcacttcaaaagtgtcgcaacatcgagctcatttccagggggaagaaatcgctcag 1800
Qy 1801	caaggacacacacacgaattgatcgaattcggctgtcattacgcgcactgtgtcgta 1860
Db 1801	caaggacacacacacgaattgatcgaattcggctgtcattacgcgcactgtgtcgta 1860
Qy 1861	caggacccctggggtgacgacaaacaaagaaatgagaagccctgcacgaaagggcagca 1920
Db 1861	caggacccctggggtgacgacaaacaaagaaatgagaagccctgcacgaaagggcagca 1920
Qy 1921	cgagaagctgctgtggaacgcgcgcacttggagcactcaccacactgcacacatcaca 1980
Db 1921	cgagaagctgctgtggaacgcgcgcacttggagcactcaccacactgcacacatcaca 1980
Qy 1981	gctggagacactggagaagcgaaggtgcgggtcggaacttgggactactcgtcctctaaa 2040
Db 1981	gctggagacactggagaagcgaaggtgcgggtcggaacttgggactactcgtcctctaaa 2040
Qy 2041	tgcactcactaatcgttctacgaaacaaataatctctactggtggtcttctgttgcaca 2100
Db 2041	tgcactcactaatcgttctacgaaacaaataatctctactggtggtcttctgttgcaca 2100
Qy 2101	ataacggtctgatacgcgcccaatttccaggaacagcccttctgttctgagatg 2160
Db 2101	ataacggtctgatacgcgcccaatttccaggaacagcccttctgttctgagatg 2160

```
QY 2161 ctcaactgtcttcgaagttgagtgctcgtcgagcaggaacgagcagactttttatctg 2220
Db 2161 ctcaactgtcttcgaagttgagtgctcgtcgagcaggaacgagcagactttttatctg 2220
QY 2221 atgtttcttctgctggtctagaaatctagtaggatatattcaacgattggtcgacatgc 2280
Db 2221 atgtttcttctgctggtctagaaatctagtaggatatattcaacgattggtcgacatgc 2280
QY 2281 aacgtttattcaacgattgctcaccatcgctatcaagcgcgaatttcccaacgagtagctg 2340
Db 2281 aacgtttattcaacgattgctcaccatcgctatcaagcgcgaatttcccaacgagtagctg 2340
QY 2341 gatcaagacatcgaaactcttcgacatcccgagcaaatcttctggtctctcaacatcgaa 2400
Db 2341 gatcaagacatcgaaactcttcgacatcccgagcaaatcttctggtctctcaacatcgaa 2400
QY 2401 ctgtcagctctaccacacgagttgcgaaggttgatatacgaacaaattttctcatttatct 2460
Db 2401 ctgtcagctctaccacacgagttgcgaaggttgatatacgaacaaattttctcatttatct 2460
QY 2461 gttgtcgtgtaacatcgctctcgagcagtgctctaccacttagcctatggtatggaactggg 2520
Db 2461 gttgtcgtgtaacatcgctctcgagcagtgctctaccacttagcctatggtatggaactggg 2520
QY 2521 ccttgagtggtgtgtgtgtgacgtctccacctgttttggctggtctacctcagaattcgg 2580
Db 2521 ccttgagtggtgtgtgtgtgacgtctccacctgttttggctggtctacctcagaattcgg 2580
QY 2581 tctagagcgaagctagaagccggaactcggcaactttgcagaagaagtgtggtggtctgc 2640
Db 2581 tctagagcgaagctagaagccggaactcggcaactttgcagaagaagtgtggtggtctgc 2640
QY 2641 aagcgaagcagttaccggtgacccgagccgtctcatctttgactctcgaaaggcactgtctct 2700
Db 2641 aagcgaagcagttaccggtgacccgagccgtctcatctttgactctcgaaaggcactgtctct 2700
QY 2701 ccaacagtaactcgacatgttgcaaggtcttgctcaagatctcaagacttttggttt 2760
Db 2701 ccaacagtaactcgacatgttgcaaggtcttgctcaagatctcaagacttttggttt 2760
QY 2761 ggaagatgttttgggttttcaactgtccacagtcagatgagtttttggctattgcccggga 2820
Db 2761 ggaagatgttttgggttttcaactgtccacagtcagatgagtttttggctattgcccggga 2820
QY 2821 tttgtattgagtcgataaattggttcaggtgagtcacacacactcaatttttatca 2880
Db 2821 tttgtattgagtcgataaattggttcaggtgagtcacacacactcaatttttatca 2880
QY 2881 tcttggtggtgttttgggttgcggttccagcagccagtcagttattggtctactccag 2940
Db 2881 tcttggtggtgttttgggttgcggttccagcagccagtcagttattggtctactccag 2940
QY 2941 agttttaccaggtcgtggtcgtggaactatctctgtggtcgtcgacattgaagcgg 3000
Db 2941 agttttaccaggtcgtggtcgtggaactatctctgtggtcgtcgacattgaagcgg 3000
QY 3001 accatcccgaaacgggaggaacaaagaaagggcccgagtcggttggtggtcgtcgac 3060
Db 3001 accatcccgaaacgggaggaacaaagaaagggcccgagtcggttggtggtcgtcgac 3060
QY 3061 ctcgagagcatgaattcaggtatcgtaacgtgattcggtcgagttctcgcggttt 3120
Db 3061 ctcgagagcatgaattcaggtatcgtaacgtgattcggtcgagttctcgcggttt 3120
QY 3121 tccatgacaatcgagcaggaacattttagcttatgtgtggcgcttctggtgtggcaag 3180
Db 3121 tccatgacaatcgagcaggaacattttagcttatgtgtggcgcttctggtgtggcaag 3180
QY 3181 tcaacgttgatcgctttgtcggaacgattctacgaccccgacactcgggccgaatttcattt 3240
Db 3181 tcaacgttgatcgctttgtcggaacgattctacgaccccgacactcgggccgaatttcattt 3240
QY 3241 gcaacagagaattatgcagaaatgtcgccgcttctgtaccgcgccatattgcttttggtc 3300
```

```
Db 3241 gcacacgagaattatgcagaaatgtcgccgcttctgtaccgcgccatattgcttttggtc 3300
QY 3301 caacaggaacacacactttaccaggtcctgttcggaagaatgtacgttgcgcctcgaa 3360
Db 3301 caacaggaacacacactttaccaggtcctgttcggaagaatgtacgttgcgcctcgaa 3360
QY 3361 gcggaattatcagaagagctttgtcaaggagcgccttcccgcaagccaatgctttggatt 3420
Db 3361 gcggaattatcagaagagctttgtcaaggagcgccttcccgcaagccaatgctttggatt 3420
QY 3421 ttgtcatctctttaccagaaagctttgaaacgcttgcggtcacaacgagggatgcagttc 3480
Db 3421 ttgtcatctctttaccagaaagctttgaaacgcttgcggtcacaacgagggatgcagttc 3480
QY 3481 tccggcggaacacagcagcagatcccatcgcaagagcattgattcgaatccaagctg 3540
Db 3481 tccggcggaacacagcagcagatcccatcgcaagagcattgattcgaatccaagctg 3540
QY 3541 ttgtacttgcaagagcagcagcctcgacacgcaatcggaacgctgtgttcaagct 3600
Db 3541 ttgtacttgcaagagcagcagcctcgacacgcaatcggaacgctgtgttcaagct 3600
QY 3601 gccctcgatgaggtatccacagcgcgaacacaaatagcagtgccgacgactttccact 3660
Db 3601 gccctcgatgaggtatccacagcgcgaacacaaatagcagtgccgacgactttccact 3660
QY 3661 attcgaagattgatttatttttttttttttttttttttttttttttttttttttttttt 3720
Db 3661 attcgaagattgatttatttttttttttttttttttttttttttttttttttttttttt 3720
QY 3721 cagcggcaactacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3780
Db 3721 cagcggcaactacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3780
QY 3781 gaccaagcatga 3792
Db 3781 gaccaagcatga 3792
```

```
RESULT 10
US-09-882-694A-10
; Sequence 10, Application US/09882694A
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; TITLE OF INVENTION: Detoxification
; FILE REFERENCE: 5718-111
; CURRENT APPLICATION NUMBER: US/09/882.694A
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 3792
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3792)
; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA
US-09-882-694A-10
```

Query Match 100.0%; Score 3791.2; DB 33; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Ds 2161 ctcaactgtctcagttgagttggtctcagtcggtcagcagagagggcagactttttattctg 2220  
Qy 2221 atgtctcttctgctcgtctcaggaatctagtaggatttctcaggtattggtcgtgacatgc 2280  
Ds 2221 atgtctcttctgctcgtctcaggaatctagtaggatttctcaggtattggtcgtgacatgc 2280  
Qy 2281 aagcttatttccaaagtgtcaccatcgtctatcaacccgcaatgttccaaacagtagctg 2340  
Ds 2281 aagcttatttccaaagtgtcaccatcgtctatcaacccgcaatgttccaaacagtagctg 2340  
Qy 2341 gatcaagacatgaactcctcagacatcccgagcaaatctctggtcgtctcacaatcgcaa 2400  
Ds 2341 gatcaagacatgaactcctcagacatcccgagcaaatctctggtcgtctcacaatcgcaa 2400  
Qy 2401 ctgtcagctctaccacagctgcaagattgatatcaagaatttctcatttattatc 2460  
Ds 2401 ctgtcagctctaccacagctgcaagattgatatcaagaatttctcatttattatc 2460  
Qy 2461 gttctcgttccacatctgtctcagcagtgctctaccactagcctatggttggaactggg 2520  
Ds 2461 gttctcgttccacatctgtctcagcagtgctctaccactagcctatggttggaactggg 2520  
Qy 2521 cctggtggtgtgtgtgtgacttccaccctgcttttggctggtctacatcgaaattcg 2580  
Ds 2521 cctggtggtgtgtgtgtgacttccaccctgcttttggctggtctacatcgaaattcg 2580  
Qy 2581 tctagagacagctagaaagcggaaactcggcaaatcttgcagaaagtctggtctgc 2640  
Ds 2581 tctagagacagctagaaagcggaaactcggcaaatcttgcagaaagtctggtctgc 2640  
Qy 2641 aagcgaagcagttacccgagatccgacccgtctcatttctgactcgaagccatgtct 2700  
Ds 2641 aagcgaagcagttacccgagatccgacccgtctcatttctgactcgaagccatgtct 2700  
Qy 2701 ccacagtagctcggacatgttgagcaaggtcttctgaagatcccaaaagcttttggtt 2760  
Ds 2701 ccacagtagctcggacatgttgagcaaggtcttctgaagatcccaaaagcttttggtt 2760  
Qy 2761 ggcagatgtttgttttcaactgtcagctcagtgaggttttggctatttgcctggga 2820  
Ds 2761 ggcagatgtttgttttcaactgtcagctcagtgaggttttggctatttgcctggga 2820  
Qy 2821 ttttctatgagctgataatctgctcaggtgagtagacacaaactcaattttatata 2880  
Ds 2821 ttttctatgagctgataatctgctcaggtgagtagacacaaactcaattttatata 2880  
Qy 2881 tctctggtggcgttttgtttgctggtcccaagcagccaggtatttggcttactccag 2940  
Ds 2881 tctctggtggcgttttgtttgctggtcccaagcagccaggtatttggcttactccag 2940  
Qy 2941 agttttaccagctcgtctgctcgaactatctcttgcgtcggagacattgaaagcgg 3000  
Ds 2941 agttttaccagctcgtctgctcgaactatctcttgcgtcggagacattgaaagcgg 3000  
Qy 3001 accatccgcaaacgagagagaaacaaagggccagtggtgagatccctctgcag 3060  
Ds 3001 accatccgcaaacgagagagaaacaaagggccagtggtgagatccctctgcag 3060  
Qy 3061 ctgagagacattgaattcaggtatcgttcaacgtgatttgcgtcaggttcccggtgtt 3120  
Ds 3061 ctgagagacattgaattcaggtatcgttcaacgtgatttgcgtcaggttcccggtgtt 3120  
Qy 3121 tccatgacaatcgagccagcaaatgtgacttattgagcgtctgagcgtctggtgcaag 3180  
Ds 3121 tccatgacaatcgagccagcaaatgtgacttattgagcgtctgagcgtctggtgcaag 3180  
Qy 3181 tcaacgtgtgattcgttctcgaagcagattctacgacccgacccggtggcgaatttcat 3240  
Ds 3181 tcaacgtgtgattcgttctcgaagcagattctacgacccgacccggtggcgaatttcat 3240  
Qy 3241 gcaacagagaatttgcagaaatgtgcgcgctgtgtacccggccatattgtcttggtc 3300  
Ds 3241 gcaacagagaatttgcagaaatgtgcgcgctgtgtacccggccatattgtcttggtc 3300

Qy 3301 caacagagaacccacacactttaccagaagctcgtttccgagagaatgtgacgttggccctcgaa 3360  
Ds 3301 caacagagaacccacacactttaccagaagctcgtttccgagagaatgtgacgttggccctcgaa 3360  
Qy 3361 gccgaattattcaagaagattttgtaagaagcgtcttccgcaagggcaatgctttggatt 3420  
Ds 3361 gccgaattattcaagaagattttgtaagaagcgtcttccgcaagggcaatgctttggatt 3420  
Qy 3421 ttgtcatctctttaccagaagctttgaaacgcttgcgttgcgttcaacgagggatcgattc 3480  
Ds 3421 ttgtcatctctttaccagaagctttgaaacgcttgcgttgcgttcaacgagggatcgattc 3480  
Qy 3481 tccggtggggcaacgacagcggatcgccatcgcaagagcatgttgcgaatcccaagctg 3540  
Ds 3481 tccggtggggcaacgacagcggatcgccatcgcaagagcatgttgcgaatcccaagctg 3540  
Qy 3541 ttgtactttgcaagaagcagctcagccctcgacgaatcggaacgtctgttcaagct 3600  
Ds 3541 ttgtactttgcaagaagcagctcagccctcgacgaatcggaacgtctgttcaagct 3600  
Qy 3601 gccctgatgagcatccacagcagcgaacgacaaatagcagtggtggccagcacttccact 3660  
Ds 3601 gccctgatgagcatccacagcagcgaacgacaaatagcagtggtggccagcacttccact 3660  
Qy 3661 attcggaattgtgatttattttgtttgttgcacgggagaaatcgccgaacgggcaact 3720  
Ds 3661 attcggaattgtgatttattttgtttgttgcacgggagaaatcgccgaacgggcaact 3720  
Qy 3721 caacggaaactcaacagcactgagaggaagataattacgagatgtgttgcacaatttta 3780  
Ds 3721 caacggaaactcaacagcactgagaggaagataattacgagatgtgttgcacaatttta 3780  
Qy 3781 gaccaagcatga 3792  
Ds 3781 gaccaagcatga 3792

RESULT 11  
US-09-351-224-9  
; Sequence 9, Application US/09351224  
; GENERAL INFORMATION:  
; APPLICANT: Duwick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351,224  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 3999  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: p-glycoprotein, with introns  
US-09-351-224-9

Query Match 91.1%; Score 3454.8; DB 17; Length 3999;  
Best Local Similarity 96.1%; Pred. No. 0;  
Matches 3658; Conservative 4; Mismatches 126; Indels 20; Gaps 11;

Qy 1 atggcagatgaatcggagaaacctcgacaaacccaagatggcagtgagctcctcacac 60  
Ds 20 atggcagatgaatcggagaaacctcgacaaacccaagatggcagtgagctcctcacac 79  
Qy 61 cctccccagaaagaaacccgaagcagattttcagactatctcagaaatcttcagatat 120  
Ds 80 cctccccagaaagaaacccgaagcagattttcagactatctcagaaatcttcagatat 139  
Qy 121 gccgacaaaatcagactggactcctcaatgtcatcgctcactctgcgcacatcgatccggg 180  
Ds 140 gccgacaaaatcagactggactcctcaatgtcatcgctcactctgcgcacatcgatccggg 199











```

|||||
Db 2418 tcaatcgcaactgacgtctaccacgcagttgcaggagttgatacaacaattctt 2477
QY ctcattatattcgttcggtcaaatcgtctcgagcagtgctctaccactagacctatgg 2508
|||||
Db 2478 ctcattttatcgttcgtacaaatcctctcgagcagtgctctagcactagacctatgg 2537
QY atggaatcaggcctcagtggtgtgttttggtgcactctccacctctgtttgagtgcta 2568
Db 2538 atggaatcaggcctcagtggtgtgttttggtgcactctccacctctgtttgagtgcta 2597
QY cctcaaatctgtctagacgaagaactagacgcaggaactcggcaaatcttgacagaag 2628
Db 2598 cctcaaatctgtctgagcgaagctagagccggaactcggcaaatcttgacagaag 2657
QY tctctgggttcgaaggaagcagttaccgcgatacccgagccctctcattcttgactctga 2688
Db 2658 tctctgggttcgaaggaagcagttaccgcgatacccgagccctctcattcttgactctga 2717
QY aggcattctctcaacagttactcggacatgttgagcaggtcttgtaagatcatccaa 2748
Db 2718 aggcattctctcaacagttactcggacatgttgagcaggtcttgtaagatcatccaa 2777
QY agcttttggtttgacagatgttttggttttcaactgttcacagtcgatgaagtttttgct 2808
Db 2778 agc-ttggtttgagacagatgttttggttttcaactgttcacagtcgatgaagtttttgct 2835
QY attgccttgagatttgat---tgacgtcgataattggtttcaagtgga-gtaacgacaa 2864
Db 2836 atggtccttggaattcttggtcatgggaagtcgactactggtctcaggtgaggtacgacaa 2895
QY actcaattttatatactctctggtggcggtttttggttcggttcgaagcagcagccagta 2924
Db 2896 actcaattttatatactctctggtggcggtttttggttcggttcgaagcagcagccagaa 2955
QY ttgtgcttactccacaggtttaccaggtctggttcggttcggttcggttcggttcggttcggt 2984
Db 2956 gcggaattactccagaggtcttaccaggtctggttcggttcggttcggttcggttcggttcggt 3015
QY gcggacattgaagccgaccatccggaagcagagagagacaaagaaagccagtgag 3044
Db 3016 gcggacattgaagccgaccatccggaagcagagagagacaaagaaagccagtgag 3075
QY tggatccctgtgacactcagagacattgaattcaaggttatcgtcaacgttatctcgctcg 3104
Db 3076 tggatccctgtgacactcagagacattgaattcaaggttatcgtcaacgttatctcgctcg 3135
QY agttctcgcgggttttccatgacaaatcgagccaggaacattttgagcttatgtgggcgc 3164
Db 3136 agttctcgcgggttttccatgacaaatcgagccaggaacattttgagcttatgtgggcgc 3195
QY ttctggtgtggcaagtcaacgttgatcggtttgtcggaacgattctacgacccgacctc 3224
Db 3196 ttctggtgtggcaagtcaacgttgatcggtttgtcggaacgattctacgacccgacctc 3255
QY gggcggaatttcatttgacacagaaatattgacaaatgtcgcgcgctgttacgcgg 3284
Db 3256 gggcggaatttcatttgacacagaaatattgacaaatgtcgcgcgctgttacgcgg 3315
QY ccatatgtctttgttccaaacagaaacccacatttaccagaggtccgttcgcgagaaatgt 3344
Db 3316 ccatatgtctttgttccaaacagaaacccacatttaccagaggtccgttcgcgagaaatgt 3375
QY gacgttgccctcgaagccgaatttatcagaagagctttgtcagaagcgtctcccgcaag 3404
Db 3376 gacgttgccctcgaagccgaatttatcagaagagctttgtcagaagcgtctcccgcaag 3435
QY gccaatgtttggtattttgcatctcttaccagaaggtttgaaacgctttcggtctca 3464
Db 3436 gccaatgtttggtattttgcatctcttaccagaaggtttgaaacgctttcggtctca 3495
QY acgagagatcagttctcggcggaacacgacgagatcccatccagagacattgat 3524
|||||
```

```

Db 3496 acgagggatgcagttctccggcgggcaacgacagcggatcccatcgcaagagcattgat 3555
QY tggaaatcgaagctgttctacttgacgaagcagcgtcagccctcgacacgcaatcgaa 3584
|||||
Db 3556 tggaaatcgaagctgttctacttgacgaagcagcgtcagccctcgacacgcaatcgaa 3615
QY acgtctggttcaagctgctcctcgatgagcattccacgagccgacgaacgaatagcagtgcc 3644
Db 3616 acgtctggttcaagctgctcctcgatgagcattccacgagccgacgaacgaatagcagtgcc 3675
QY gcaccgactttccactattcggaaattgtatgttttttttttttttttttttttttttttttt 3704
Db 3676 gcaccgactttccactattcggaaattgtatgttttttttttttttttttttttttttttttt 3735
QY ggcgaacggggaactcagcggaactacacgactgagaggaagatatattacagatgtcg 3764
Db 3736 ggcgaacggggaactcagcggaactacacgactgagaggaagatatattacagatgtcg 3795
QY ttggcacaactcttttagaccaagcatga 3792
Db 3796 ttggcacaactcttttagaccaagcatga 3823

RESULT 13
US-09-351-823-9
; Sequence 9, Application US/09351823
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification
; FILE REFERENCE: 5718-56
; CURRENT APPLICATION NUMBER: US/09/351,823
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; OTHER INFORMATION: p-glycoprotein, with introns
US-09-351-823-9

Query Match 91.1%; Score 3454.8; DB 17; Length 3999;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3658; Conservative 4; Mismatches 126; Indels 20; Gaps 11;

QY 1 atggcagatgaatcggagaaacctcgaccaaaccaagatggcagtgagtcgtctcacac 60
|||||
Db 20 atggcagatgaatcggagaaacctcgaccaaaccaagatggcagtgagtcgtctcacac 79
QY 61 cctccccagaaaagaaaccgagggcagttatttcagactatctacgaattcttcagatat 120
Db 80 cctccccagaaaagaaaccgagggcagttatttcagactatctacgaattcttcagatat 139
QY 121 gccgacaatacagactgagctcgaatgtcgaatgtcgcgcgtcctcctcgccacgcgacgg 180
|||||
Db 140 gccgacaatacagactgagctcgaatgtcgaatgtcgcgcgtcctcctcgccacgcgacgg 199
QY 181 gttccctctctctgagtgatcgtatctctcgtagcttcacaaacgaagttcaacaattac 240
Db 200 gttccctctctctgagtgatcgtatctctcgtagcttcacaaacgaagttcaacaattac 259
QY 241 aattcggcgagcggagttccttgaaagcgttcaagcgttcaagcgttcaagcgttcaagcgtt 300
Db 260 aattcggcgagcggagttccttgaaagcgttcaagcgttcaagcgttcaagcgttcaagcgtt 319
QY 301 ttgcgtacctctttatttggaagttttgctcagctacgtttccacggcgttcgacattacc 360
Db 320 ttgcgtacctctttatttggaagttttgctcagctacgtttccacggcgttcgacattacc 379
QY 361 attcagctatcacgaacccctcgaactcttcgacagtggttcttcgaatgcaccttcgag 420
|||||
```

Dbb 380 atttcagctatacgaacacactcgaactcttctgacagtggttcttcttgaatgcacacttgccg 439  
Qy 421 caaagagctctgagcttccacaaagacagcaaatggagcaatgcgcaactcargtcaactacc 480  
Dbb 440 caaagggctcggcattctcgaaagcagcaaatggagcaatgcgcaactcargtcaactacc 499  
Qy 481 aatggcaacogtatacaaaacaggtatgcgagaaatgggttttaacogtgcagggcaactt 540  
Dbb 500 aatggcaacogtatacaaaacaggtatgcgagaaatgggttttaacogtgcagggcaactt 559  
Qy 541 tcaatgtttttctgacatttggtgcgttttggcgttcagtgagtgagctagcttttaac 600  
Dbb 560 tcaatgtttttctgacatttggtgcgttttggcgttcagtgagtgagctagcttttaac 619  
Qy 601 accatgtccgtacccctgcatttctcgttgcacggcatctcatagcaaatgtatgcc 660  
Dbb 620 accatgtccgtacccctgcatttctcgttgcacggcatctcatagcaaatgtatgcc 679  
Qy 661 gctcagagggcaggatccacagatctactcaacgcgcgtctgctcctcgcagaagaagtc 720  
Dbb 680 gctcagagggcaggatccacagatctactcaacgcgcgtctgctcctcgcagaagaagtc 739  
Qy 721 ttaataccatccggacagttccatgtttctacgccagagaagaatggtcgaaaaatat 780  
Dbb 740 ttaataccatccggacagttccatgtttctacgccagagaagaatggtcgaaaaatat 799  
Qy 781 gatgtcttttgcagcaagcacaccaagaaggaagaagaatcgccaaataatggsgtc 840  
Dbb 800 gatgtcttttgcagcaagcacaccaagaaggaagaagaatcgccaaataatggsgtc 859  
Qy 841 ttgtcttcaactagtaactttgcatctacgtctatcgcaactggcctttt--gaaa 897  
Dbb 860 ttgtcttcaactagtaactttgcatctacgtctatcgcaactggccttttgggaaag 919  
Qy 898 ggttttgcagtatcagaatgcaggtccgacagttgcaagcttgaagcttacttggtc 957  
Dbb 920 gtttttgcagtatcagaatggcgggttgcgcagcttggcaagcttacttacttgcgc 979  
Qy 958 ctttcogtcaacttagcagcaactccatccaatgcctgcgtcctcaggttcaagtcgtt 1017  
Dbb 980 ttttcogtcaacttagcagcaactccatccaatgcctgcgtcctcaggttcaagtcgtt 1039  
Qy 1018 taccacgcgcgcatcttcgg--ctccgaattattcagtatcatattacaaacccacgcagct 1076  
Dbb 1040 taccacgcgcgcatcttcggcctccgaattattcagtatcatattacaaacccacgcagct 1099  
Qy 1077 cgacct--ctcgaccttttgaaagcagccagaggggtcgtcaggtcgaatttgagatc 1134  
Dbb 1100 cgaccttttcgaccttttggaaagcagccagaggggtcgtcaggtcgaatttgagatc 1159  
Qy 1135 caaaacgtgcatttgacctccacccctccgacacatctgcgaagtaactcagagattcaac 1194  
Dbb 1160 caaaacgtgcatttgacctccacccctccgacacatctgcgaagtaactcagagattcaac 1219  
Qy 1195 ttgcaatttcaagctggcgaagcagcggccctcgtcgtgtgcatcaggtagcggcaaaagc 1254  
Dbb 1220 ttgcaatttcaagctggcgaagcagcggccctcgtcgtgtgcatcaggtagcggcaaaagc 1279  
Qy 1255 acaatgtcgtcgttacttgaacggttgatctgcgccagttcggggagagatattacttgat 1314  
Dbb 1280 acaatgtcgtcgttacttgaacggttgatctgcgccagttcggggagagatattacttgat 1339  
Qy 1315 ggggttgaactgggacaatacaatgtgaaatggctggaagccgcatctcctcgttcaa 1374  
Dbb 1340 ggggttgaactgggacaatacaatgtgaaatggctggaagccgcatctcctcgttcaa 1399  
Qy 1375 caggaacgtgtgttttgcgtggcacaactctccagaacattgcgaacggtttcaatggat 1434  
Dbb 1400 caggaacgtgtgttttgcgtggcacaactctccagaacattgcgaacggtttcaatggat 1459  
Qy 1435 gagcaacagatctcgtcgtcgaaaaaacaataatggagcttgcgaagaagcttgcgaagcc 1494  
Dbb 1460 gagcaacagatctcgtcgtcgaaaaaacaataatggagcttgcgaagaagcttgcgaag-c 1518

Qy 1495 agcaatggcgagcgtgttcatttaactgagcttcggaacggttatcgagactgaagtggcgag 1554  
Dbb 1519 agcaatggcgagcgtgttcatttaactgagcttcggaacggttatcgagactgaagtggcgag 1578  
Qy 1555 cgagccggagccttgagtgaggttcaacgcgaacaaatgtgcaatcgcaacgaagtataata 1614  
Dbb 1579 cgagccggagccttgagtgaggttcaac--aagccgaatgtgcaatcgcaacgaagtataata 1637  
Qy 1615 tcggatcccaagatcctgttactctgagctgagcgccttgcagcccttgcaggaagggcgag 1674  
Dbb 1638 tcggatcccaagatcctgttactctgagctgagcgccttgcagcccttgcaggaagggcgag 1697  
Qy 1675 aagtggtccagagagccttgaaacggagtgccaaagacgcactacttctggctcaattgcc 1734  
Dbb 1698 aagtggtccagagagccttgaaacggagtgccaaagacgcactacttctggctcaattgcc 1757  
Qy 1735 caaaactagcactgtcaaaaagtctggaacatcgagctcatttcccaggggaaaaatc 1794  
Dbb 1758 caaaactagcactgtcattacgactcactataggcgcaattggcctctagatgcat 1817  
Qy 1795 g----tcgagcaagcacaac 1849  
Dbb 1818 gctcgagcgcccgccagtgtagcaatgtgtagcaatgtcggaattcggtctgtatcgcgac 1877  
Qy 1850 tgggtcgctgcagagggcctcggggtgcagcaacaaagaaacacacacacacacacacacac 1909  
Dbb 1878 tgggtcgctgcagagggcctcggggtgcagcaacaaagaaacacacacacacacacacacac 1937  
Qy 1910 aaaaagcagcagagagcgtgctggtgaacgcacggcgcacttgagcgcactcacacacacg 1969  
Dbb 1938 aaaaagcagcagagagcgtgctggtgaacgcacggcgcacttgagcgcactcacacacacg 1997  
Qy 1970 ccacatctcaagctgagagcctggagaagcggaggtgcgggtcgagacattgggactt 2029  
Dbb 1998 ccacatctcaagctgagagcctggagaagcggaggtgcgggtcgagacattgggactt 2057  
Qy 2030 cgtctcaaaatgcactcctcaatcactgtctctacgaacaaaaaaactctactggtctct 2089  
Dbb 2058 cgtctcaaaatgcactcctcaatcactgtctctacgaacaaaaaaactctactggtctct 2117  
Qy 2090 tgggtcaacaataaagcgttctgatatgctggcgccacatttcccagggacacgcttttgt 2149  
Dbb 2118 tgggtcaacaataaagcgttctgatatgctggcgccacatttcccagggacacgcttttgt 2177  
Qy 2150 ttcgagattgctcactgtctctcaggttgagtggttcagtcgagcaggaacggcgagact 2209  
Dbb 2178 ttcgagattgctcactgtctctcaggttgagtggttcagtcgagcaggaacggcgagact 2237  
Qy 2210 tttatattctgagttcttctgctgagcttaggaactctagtaggatatattcagattg 2269  
Dbb 2238 tttatagctcagcttcttctgctgagcttaggaactctagtaggatatattcagattg 2297  
Qy 2270 gctggacatgcagcttatttctcacaagtgttcaacacatgcgctatcgaagcgcgaattgtcc 2329  
Dbb 2298 gctggacatgcagcttatttctcacaagtgttcaacacatgcgctatcgaagcgcgaattgtcc 2357  
Qy 2330 aacgagctactgcagac 2389  
Dbb 2358 aacgagctactgcagac 2417  
Qy 2390 tccatcgcaactgtcagctctacccacgcagcttgcaagaggtatcatcagcaaat-tt 2448  
Dbb 2418 tccatcgcaactgtcagctctacccacgcagcttgcaagaggtatcatcagcaaat-tt 2477  
Qy 2449 ctcatattatggttctcgttcaacatcgtctcagcagcaggtctctacacacacacacacacac 2508  
Dbb 2478 ctcatattctcgtctcgttcaac 2537  
Qy 2509 atggaaactggcctcgttgggtgtgttctgggtgcacttccacccctgcttttggctggcaca 2568  
Dbb 2538 atggaaactggcctcgttgggtgtgttctgggtgcacttccacccctgcttttggctggcaca 2597



QY 541 tcaatgttcttctgtgattgttggtgctgttggtgctgtcagtggaagctagctttaacg 600  
DB 560 tcaatgttcttctgtgattgttggtgctgttggtgctgtcagtggaagctagctttaacg 619  
QY 601 accatgtcgttcacccctgtgctgttctgtgttcacccgcatctgcatagcaatgtatgcc 660  
DB 620 accatgtcgttcacccctgtgctgttctgtgttcacccgcatctgcatagcaatgtatgcc 679  
QY 661 gctcgggggacagatcacagagatctactacgcgcgcgcgtgctcctcgcagagaagctc 720  
DB 680 gctcaggaggacagatcacagagatctactacgcgcgcgcgtgctcctcgcagagaagctc 739  
QY 721 ttatcatcatcccgagacgtccatctgttctacgcccagagaaatgttcgaaaaatat 780  
DB 740 ttatcatcatcccgagacgtccatctgttctacgcccagagaaatgttcgaaaaatat 799  
QY 781 gatgtcttcttgcagcagacacccaagagggaagaaataatcccaaaaataatggsgtc 840  
DB 800 gatgtcttcttgcagcagacacccaagagggaagaaataatcccaaaaataatggsgtc 859  
QY 841 ttgttctcaactgagttacttttgcatttcaactgctgctatgcactggcctttt--gaaa 897  
DB 860 ttgttctcaactgagttacttttgcatttcaactgctgctatgcactggccttttgggaaa 919  
QY 898 gatttctcatgtatcagaatggcagaggttgcagcgttggcagagtttctactgttctc 957  
DB 920 gatttctcatgtatcagaatggcagaggttgcagcgttggcagagtttctactgttctc 979  
QY 958 ctcttcgctacgttagcagcagctcatctcaatgttgcgccttcaggttcaagtcgtt 1017  
DB 980 ttctcgtcaccttagcagcagctcatctcaatgttgcgccttcaggttcaagtcgtt 1039  
QY 1018 taccagccgcatcttcg--ctccaaattattcagatcatctacaaacccacgcagct 1076  
DB 1040 taccagccgcatcttcgcttcggaattattcagatcatctacaaacccacgcagct 1099  
QY 1077 cgaacct--ctcgcacctcttgaaagcagcaggggtgctgctaggttcaaatgtagatc 1134  
DB 1100 cgaacctctcgcagcccttttggaaagcagcaggggtgctgctaggttcaaatgtagatc 1159  
QY 1135 caaaaccttgcatctgctacccctccgaccatctgcaccaagttacttcgagatttcaac 1194  
DB 1160 caaaaccttgcatctgctacccctccgaccatctgcaccaagttacttcgagatttcaac 1219  
QY 1195 ttgacaattccagctggcagacgagcgcctctgctgctgctgctgctgctgctgctgct 1254  
DB 1220 ttgacaattccagctggcagacgagcgcctctgctgctgctgctgctgctgctgctgct 1279  
QY 1255 acaatgctcgttacttgacgttggtatctgcacagttcggggagagatattacttgat 1314  
DB 1280 acaatgctcgttacttgacgttggtatctgcacagttcggggagagatattacttgat 1339  
QY 1315 ggggtggaaactgggacaatacatgtgaatggctggaagcgcgcatctgcctcgttcaa 1374  
DB 1340 ggggtggaaactgggacaatacatgtgaatggctggaagcgcgcatctgcctcgttcaa 1399  
QY 1375 cagggaacctgtgtgttctggtgcaacaattctccgaaacatttgcacacgggttctatgat 1434  
DB 1400 cagggaacctgtgtgttctggtgcaacaattctccgaaacatttgcacacgggttctatgat 1459  
QY 1435 gagcaacagatctgctcgcgaaaaaacaataatggagctgttgcacaaagcttgcacagcc 1494  
DB 1460 gagcaacagatctgctcgcgaaaaaacaataatggagctgttgcacaaagcttgcacag-c 1518  
QY 1495 agcaatgcgcagctgttcaatgatgcttccgaacggtttagagactgaagtggcgag 1554  
DB 1519 agcaatgcgcagctgttcaatgatgcttccgaacggtttagagactgaagtggcgag 1578  
QY 1555 cgaacggagccttgagtgaggttcaacacacgaatgaatgaatgcgcaggaagtatcata 1614  
DB 1579 cgaacggagccttgagtgaggttcaac-aagccgaattgcaatgcgcaggaagtatcata 1637

QY 1615 tccgattcccaagatctctgttactcagatgaagctacacagcccttgaccacgaagcagag 1674  
DB 1638 tccgattcccaagatctctgttactcagatgaagctacacagcccttgaccacgaagcagag 1697  
QY 1675 aaagtggctcagaggagccttgaaacagtgctcacaagaccgacactacttttggtatgcc 1734  
DB 1698 aaagtggctcagaggagccttgaaacagtgctcacaagaccgacactacttttggtatgcc 1757  
QY 1735 cacaacttagccactgtcaaaagtgtggaacatcgacgtcatttcccgagggaatac 1794  
DB 1758 cacaacttagccactgtcatacagactcatttagggcgcaattgggcccctctagatgat 1817  
QY 1795 g-----tcgagcaaggcacacaccacgaattgtcgaattcgtcgtctacacgcgcac 1849  
DB 1818 gctcgaagcgcgcgcagtgagcaaatgtgagcaaatcgtcgtctacacgcgcac 1877  
QY 1850 tggctcgtgcacaggaacctcggggctgagcaacaacaagaacatgagagaccctgcacg 1909  
DB 1878 tggctcgtgcacaggaacctcggggctgagcaacaacaagaacatgagagaccctgcacg 1937  
QY 1910 aaaaggcagcacgagaagctgctggtgaacgaccgacacttgagcgcactcacaccactg 1969  
DB 1938 aaaaggcagcacgagaagctgctggtgaacgaccgacacttgagcgcactcacaccactg 1987  
QY 1970 ccacatctcaagctggagacactggagagcggagagtgccggtcgggactttgggtact 2039  
DB 1998 ccacatctcaagctggagacactggagagcggagagtgccggtcgggactttgggtact 2057  
QY 2030 cgcctctcaaatgcatcctaatactgttctacgaaacaaataatctctactcgggtctct 2089  
DB 2058 cgcctctcaaatgcatcctaatactgttctacgaaacaaataatctctactcgggtctct 2117  
QY 2090 tcttgcacaacaataagcttctgatatgcgcgcacacattccaggaacacacccctttgt 2149  
DB 2118 tcttgcacaacaataagcttctgatatgcgcgcacacattccaggaacacacccctttgt 2177  
QY 2150 tctcgaagatgctcactgtcttcgaattgagtggtcagcagacaggaagcagact 2209  
DB 2178 tctcgaagatgctcactgtcttcgaattgagtggtcagcagacaggaagcagact 2237  
QY 2210 tttatatctcgaagttcttctgctgctcaggaatacttagtaggattctcagattg 2269  
DB 2238 tttatagctgatttcttctgctgctcaggaatacttagtaggattctcagattg 2297  
QY 2270 gctggacatgcaacgcttatttcaacagttgttccaccatcgcctatcagcgcgaatttcc 2329  
DB 2298 gctggacatgcaacgcttatttcaacagttgttccaccatcgcctatcagcgcgaatttcc 2357  
QY 2330 aacgagttactggatcagacatcgaactcctcgaactcccgagacaaatctctggtctc 2389  
DB 2358 aacgagttactggatcagacatcgaactcctcgaactcccgagacaaatctctggtctc 2417  
QY 2390 tcaatcgcacactgtcagcttaccacgcagttgcaagagttgatatcagcaaat-tt 2448  
DB 2418 tcaatcgcacactgtcagcttaccacgcagttgcaagagttgatatcagcaaatctct 2477  
QY 2449 ctcattttatcgttctggtgcaacatcgtctcgaagcagtgcttaccacagactagcctagg 2508  
DB 2478 ctcattttatcgttctggtgcaacatcgtctcgaagcagtgcttaccacagactagcctagg 2537  
QY 2509 atggaactggcctggt 2568  
DB 2538 atggaactggcctggt 2597  
QY 2569 cctcagaatctgtatgagacgaagctagaaacccggaacactcgcgaactttgcagaag 2628  
DB 2598 cctcagaatctgtatgagacgaagctagaaacccggaacactcgcgaactttgcagaag 2657  
QY 2629 tgcgtggcttgcagcagcaggttaccgcgattccggacgttctctatcttctgactcga 2688  
DB 2658 tgcgtggcttgcagcagcaggttaccgcgattccggacgttctctatcttctgactcga 2717  
QY 2689 aggcattgtcttccaaacagttactcgcagatgttgagcaaggcttctgtcgaagatcatcaa 2748





||||| 560 tcaatgtctcttctgcatcttggtcgctttggtctcagtggaactagctttaac 619  
QY 601 accatgtctgcctcatccctgcactttctctggtccacggcgcattgcatagcaaatgagtc 660  
Db 620 accatgtctgcctcatccctgcactttctctggtccacggcgcattgcatagcaaatgagtc 679  
QY 661 gctcaggaggcaggatcacaggatctactcagcgccgtgctctgcagagaagtc 720  
Db 680 gctcaggaggcaggatcacaggatctactcagcgccgtgctctgcagagaagtc 739  
QY 721 ttatcatcatcccgagcagtcctctctcagcccgagagaagaatggtcgaaaaatat 780  
Db 740 ttatcatcatcccgagcagtcctctctcagcccgagagaagaatggtcgaaaaatat 799  
QY 781 gatgtctctttgcagcagcacaccagagagaagaagaatcccaaatcaatggagtc 840  
Db 800 gatgtctctttgcagcagcacaccagagaggaagaagaatcccaaatcaatggagtc 859  
QY 841 ttgtctcaactgagctctttgcattctacgtgctatgcgaactggccctttt---gaaa 897  
Db 860 ttgtctcaactgagctctttgcattctacgtgctatgcgaactggcccttttgggaaag 919  
QY 898 ggtttctgcattgcatcagaatggcgaggttgcgcagcttggcaaaagtcttactgttgc 957  
Db 920 gttttctgcattgcatcagaatggcgaggttgcgcagcttggcaaaagtcttactgttgc 979  
QY 958 cttttctgcattgcatcagaatggcgaggttgcgcagcttggcaaaagtcttactgttgc 1017  
Db 980 tttctgcatttgcagcagcagtcctcatctcaatgcttgcgcagcttgcaggttgcagtc 1039  
QY 1018 taccacgcgcattctcgg-cctcgaattattcagttcattgacaaacacacagagct 1076  
Db 1040 taccacgcgcattctcggcctccgaattattcagttcattgacaaacacacagagct 1099  
QY 1077 cgaacct---ctcgaacctcttgaaagcgcaagagggctgctaggttcaaatlgagatc 1134  
Db 1100 cgaacctctcgaacctcttgaaagcgcaagagggctgctaggttcaaatlgagatc 1159  
QY 1135 caaaacctggcatttgcctaccctcccgaccatctgcgaagtacttcgagatttcaac 1194  
Db 1160 caaaacctggcatttgcctaccctcccgaccatctgcgaagtacttcgagatttcaac 1219  
QY 1195 ttgacaaattccagctggcaagacagcgccctcgtctggtgcatcaggtagcgcaaaagc 1254  
Db 1220 ttgacaaattccagctggcaagacagcgccctcgtctggtgcatcaggtagcgcaaaagc 1279  
QY 1255 acaatgtcggcttacttgaaactggtatctgcgaacttcggaagagataattactgct 1314  
Db 1280 acaatgtcggcttacttgaaactggtatctgcgaacttcggaagagataattactgct 1339  
QY 1315 ggttggaactcgggaacaatacaatgtgaaatggctgagaagcgcattcgcctcgttcaa 1374  
Db 1340 ggttggaactcgggaacaatacaatgtgaaatggctgagaagcgcattcgcctcgttcaa 1399  
QY 1375 caggaaacctgtgttcttgccacaaattctccagaacattgcgaacgggttctcatggat 1434  
Db 1400 caggaaacctgtgttcttgccacaaattctccagaacattgcgaacgggttctcatggat 1459  
QY 1435 gagcaacgagatcgtcctcgcgaaaaaaaatggagcttgtgcaaaaaagcttgcgaagcc 1494  
Db 1460 gagcaacgagatcgtcctcgcgaaaaaaaatggagcttgtgcaaaaaagcttgcgaag-c 1518  
QY 1495 agcaatggcagctgtctcattatgacttcgaacggttatgagactgaagtggcgag 1554  
Db 1519 agcaatggcagctgtctcattatgacttcgaacggttatgagactgaagtggcgag 1578  
QY 1555 cgaacggagccttgagtgagtgcaacaaacgaacgaatgaatgcgaagtatcata 1614  
Db 1579 cgaacggagccttgagtgagtgcaac--aagccgaatgtgcaacgcgaagtatcata 1637  
QY 1615 tggatcccaagatcccttactcgtatgagctaccagcgcccttgaccggaagcgag 1674  
|||||

Db 1638 tcggtatcccaagatcctgttactcgtatgaagctaccagcgcccttgaccgagcgag 1697  
QY 1675 aaatgtctcagagagccttgaaacagagtgctcaaaagaccgcactactcttggctatgccc 1734  
Db 1698 aaatgtctcagagagccttgaaacagagtgctcaaaagaccgcactactcttggctatgccc 1757  
QY 1735 cacaactagccactgtcacaagagtgctggcaacatcgcagtcatttcccaggggaaaaatc 1794  
Db 1758 cacaactagccactgtcacaagagtgctgaactcactattagggaattgggccccttagatgcat 1817  
QY 1795 g-----tcgagcagagcgcacacacagaaattgataatggctgctcattacccgcac 1849  
Db 1818 gctcgagcgcccgccagtgtagcgaattgtgcgaatttgcggttgcattacccgcac 1877  
QY 1850 tgggtcgtgcacagcctcgggctgcagcgaacacagacacatgagaagaccctgcacg 1909  
Db 1878 tgggtcgtgcacagcctcgggctgcagcgaacacagacacatgagaagaccctgcacg 1937  
QY 1910 aaaaggcagcgcagagagctgctggtcgaacgcggcacttgagcgcaactcacaccactg 1969  
Db 1938 aaaaggcagcgcagagagctgctggtcgaacgcggcacttgagcgcaactcacaccactg 1997  
QY 1970 ccacatctcaagctggagacccggagagcggaaggtgcccgtcgggacttgggctact 2029  
Db 1998 ccacatctcaagctggagacccggagagcggaaggtgcccgtcgggacttgggctact 2057  
QY 2030 cgtcctcaaaatgcattcctaaatcattctctacgaacaaaaaatctctactggtgctct 2089  
Db 2058 cgtcctcaaaatgcattcctaaatcattctctacgaacaaaaaatctctactggtgctct 2117  
QY 2090 tgggtcgaacaaatgaacggcttgcatacgcggccacatcttccaggacaccccttctgt 2149  
Db 2118 tgggtcgaacaaatgaacggcttgcatacgcggccacatcttccaggacaccccttctgt 2177  
QY 2150 tttcgaagatgctcactgtctctcgaagtgagtggtcgaagcgcaagcggaagcgagact 2209  
Db 2178 tttcgaagatgctcactgtctctcgaagtgagtggtcgaagcgcaagcggaagcgagact 2237  
QY 2210 tttatattctgagtgcttcttctgctggtcgttaggaaactcgtagtaggattctcagactg 2269  
Db 2238 tttatagctgagtgcttcttctgctggtcgttaggaaactcgtagtaggattctcagactg 2297  
QY 2270 gctggacatgcacgcttatttccaaagtggtcaccacatcgcctatcaagcgcaaatgttcc 2329  
Db 2298 gctggacatgcacgcttatttccaaagtggtcaccacatcgcctatcgcagcgcaaatgttcc 2357  
QY 2330 aacgagactcagatcagacatcgcactcctcgcacatcccgagacaaatcttctgctgctc 2389  
Db 2358 aacgagactcagatcagacatcgcactcctcgcacatcccgagacaaatcttctgctgctc 2417  
QY 2390 tcacatcgcactcagatcagatcctaccacagcagctgcgaagagttgatatcgcgaacatc-tt 2448  
Db 2418 tcacatcgcactcagatcagatcctaccacagcagctgcgaagagttgatatcgcgaacatcctt 2477  
QY 2449 ctcaatttatactgttctcgtgtcaacatcgtctcgcagcgagtgctctacccactagcctatgg 2508  
Db 2478 ctcaatttatactgttctcgtgtcaacatcctctcgcagcgagtgctctacccactagcctatgg 2537  
QY 2509 atggaactcggcctggtggtggtgttctgctcacttccaccctgcttcttggctgctga 2568  
Db 2538 atggaactcggcctggtggtggtgttctgctcacttccaccctgcttcttggctgctga 2597  
QY 2569 cctcagaatctcgtcagacgaagcagatgagaacgcgaagcctcgcgaacttgcgaacttgcagaag 2628  
Db 2598 cctcagaatctcgtcagacgaagcagatgagaacgcgaagcctcgcgaacttgcagaag 2657  
QY 2629 tgcctggtcgtcgaagcgaagcagatgagaacgcgaagcctcgcgaacttgcgaacttgcaga 2688  
Db 2658 tgcctggtcgtcgaagcgaagcagatgagaacgcgaagcctcgcgaacttgcgaacttgcaga 2717  
QY 2689 aggcacttctcccaacagctactcggacatgttgagcgaaggtcttgcgaagatcctccaa 2748  
Db 2718 aggcacttctcccaacagctactcggacatgttgagcgaaggtcttgcgaagatcctccaa 2777



Search completed: April 27, 2002, 01:38:22  
Job time: 22266 sec

QY	2749	agcttttggtttggacgattgtttttgttttttcaacttgcacagtgcagtgcagtttttggct	2800
Db	2778	agc-tttgggtttggacgagtggtttgtttctcac-tgtcacagctcagatcgagtttttggct	2835
QY	2809	attgcctcgggatttttgat---tgcagtcgataatttgcttcagggtga-gtacacacaa	2864
Db	2836	atggcctcgggaatttttgatctgggaagtgagctacttggtcttcadgtgaggtacacaca	2895
QY	2865	actcaatttatatcatcttcgtggcggtttgtttggcgcgtcccaagcagcagccagta	2924
Db	2896	actcaatttatatcatcttcgtggcggtttgtttggcgcgtcccaagcagcagccagaa	2955
QY	2925	tttggttaactcacagattttaccaaagctccggttcggctgcgaactatctcttgct	2984
Db	2956	gccgaattactccagagctttaccaaagctcggctgcgaactatctcttgct	3015
QY	2985	ggcgacattgaagcccgacacatccgcgaaaaacggagagagacaagaanaagccagtcgg	3044
Db	3016	ggcgacattgaagcccgacacatccgcgaaaacggagagagacaagaanaagccagtcgg	3075
QY	3045	tggatggcctgtcgacctcgagacattggaattcaaggtatgctcaactgattcggctcg	3104
Db	3076	tggatggcctgtcgacctcgagacattggaattcaaggtatgctcaactgattcggctcg	3135
QY	3105	agttctccgcgggtttccatgacaaatcgagccagagacaattttagcttatgtgggcgc	3164
Db	3136	agttctccgcgggtttccatgacaaatcgagccagagacaattttagcttatgtgggcgc	3195
QY	3165	tcttggtctggaagtcacagcttgatcgctttgtcggaacgattctcagccgcacct	3224
Db	3196	tcttggtctggaagtcacagcttgatcgctttgtcggaacgattctcagccgcacct	3255
QY	3225	ggcgcaatttcatttgcacacagagaattatgcagaaaatgtcgccgccttgcacccgg	3284
Db	3256	ggcgcaatttcatttgcacacagagaattatgcagaaaatgtcgccgccttgcacccgg	3315
QY	3285	ccatatgtctttgttccaaagagaccacactttaccaggctcgttcgcagaaatgt	3344
Db	3316	ccatatgtctttgttccaaagagaccacactttaccaggctcgttcgcagaaatgt	3375
QY	3345	gacgtttgccttcgaagccgaattatcagaagagctttgtcgaagcagccttcgcgaag	3404
Db	3376	gacgtttgccttcgaagccgaattatcagaagagctttgtcgaagcagccttcgcgaag	3435
QY	3405	gccatgcttggaatttgcattctcttaccagaagctttgaaacgcttcgcgcctca	3464
Db	3436	gccatgcttggaatttgcattctcttaccagaagctttgaaacgcttcgcgcctca	3495
QY	3465	acgaaggtatgaatttctccgcggccaaacgacagcgatgcgcacgaagcattgat	3524
Db	3496	acgaaggtatgaatttctccgcggccaaacgacagcgatgcgcacgaagcattgat	3555
QY	3525	tcgaataccaaagcttgtctacttgcagagcgacgttcagccctgcagacgaatcgg	3584
Db	3556	tcgaataccaaagcttgtctacttgcagagcgacgttcagccctgcagacgaatcgg	3615
QY	3585	acgtctggtttcaagtcgctcgaagaggtatccacagccgacgaataagcagtcggc	3644
Db	3616	acgtctggtttcaagtcgctcgaagaggtatccacagccgacgaataagcagtcggc	3675
QY	3645	gcacgactttccactattcgaagatttgatgtattttgtttgttcgcaacgaggaat	3704
Db	3676	gcacgactttccactattcgaagatttgatgtattttgtttgttcgcaacgaggaat	3735
QY	3705	cgccgaacgggacctcagcgggaactacaacgactgagaggagaattatcagagatgtg	3764
Db	3736	cgccgaacgggacctcagcgggaactacaacgactgagaggagaattatcagagatgtg	3795
QY	3765	tttggcacaactcttagaccaagcatga	3792
Db	3796	tttggcacaactcttagaccaagcatga	3823

-----

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 01:35:58 ; Search time 150.7 Seconds  
(without alignments)  
6180.771 Million cell updates/sec

Title: US-09-882-694A-10  
Perfect score: 3792  
Sequence: 1 atggcagatgaatcgagaa.....aatcttttagaccagcatga 3792

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues  
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723.6	19.1	3927	2	US-08-996-644-1
2	723.6	19.1	3927	3	US-09-352-552-1
3	722.2	19.0	3924	2	US-08-996-644-3
4	722.2	19.0	3924	3	US-09-352-552-3
5	487	12.8	4002	2	US-08-996-545-1
6	487	12.8	4002	2	US-08-996-545-3
7	487	12.8	4002	4	US-09-328-320-1
8	487	12.8	4002	4	US-09-328-320-3
9	450.6	11.9	4047	2	US-08-612-734B-1
10	367.4	9.7	4800	2	US-08-612-734B-3
11	362	9.5	3924	1	US-08-395-246C-1
12	351.2	9.3	4224	1	US-08-612-521-1
13	286	7.5	4233	3	US-09-120-513-1
14	286	7.5	4233	4	US-09-450-105-1
15	270.2	7.1	4646	1	US-08-181-471-2
16	270.2	7.1	4669	6	5206352-3
17	270.2	7.1	6505	2	US-08-793-610-5
18	270.2	7.1	9318	2	US-08-793-610-6
19	263.8	7.0	4669	2	US-08-752-447-1
20	259	6.8	4264	2	US-08-784-649A-1
21	259	6.8	4264	2	US-08-784-649A-5
22	255.8	6.7	4669	2	US-08-583-276-18
23	162.6	4.3	2725	1	US-08-461-823-1
24	136.4	3.6	3909	1	US-08-232-537-1
25	128.6	3.4	6143	1	US-08-612-521-3
26	111.8	2.9	28804	2	US-08-592-874-1
27	111.8	2.9	28804	3	US-09-096-942-2

c	28	111.8	2.9	28804	3	US-09-096-867-2	Sequence 2, Appli
c	29	110.4	2.9	3576	4	US-08-976-259-79	Sequence 79, Appli
c	30	105	2.8	535	4	US-08-858-207A-230	Sequence 230, Appli
c	31	104.4	2.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	32	104	2.7	7721	3	US-08-772-270A-14	Sequence 14, Appli
	33	103.6	2.7	5120	3	US-08-772-270A-6	Sequence 6, Appli
	34	103.6	2.7	8370	2	US-08-488-706-1	Sequence 1, Appli
	35	103.4	2.7	2407	2	US-08-895-522-2	Sequence 2, Appli
	36	103.4	2.7	2407	3	US-09-195-391-2	Sequence 2, Appli
	37	103.4	2.7	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	38	101.2	2.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	39	100.4	2.6	2376	1	US-08-394-880B-1	Sequence 18, Appli
	40	99.6	2.6	2244	4	US-09-061-764A-18	Sequence 189, Appli
	41	97.6	2.6	1201	3	US-08-961-083-189	Sequence 17, Appli
	42	85.2	2.2	2061	4	US-09-061-764A-17	Sequence 1, Appli
	43	80	2.1	4394	2	US-08-750-152A-1	Sequence 4, Appli
	44	76.8	2.0	1959	4	US-09-061-764A-4	Sequence 1, Appli
	45	69.6	1.8	396	2	US-09-108-837-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-996-644-1  
; Sequence 1, Application US/08996644  
; Patent No. 5945324  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrc of  
; TITLE OF INVENTION: Aspergillus Nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,644  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11765  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3927 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..3924  
US-08-996-644-1

Query Match 19.1%; Score 723.6; DB 2; Length 3927;  
Best Local Similarity 52.1%; Pred. No. 1.2e+228;





```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3924
US-09-352-552-1

```

Query Match 19.1%; Score 723.6; DB 3; Length 3927;  
Best Local Similarity 52.1%; Pred. No. 1.2e-228;  
Matches 1929; Conservative 1; Mismatches 1745; Indels 29; Gaps 13;

QY	107	gaatttcagatgacgcagaataatcgactggactctcaatgtcatcgcgcctcactcgcg	165
Db	229	GGATCTTTACATTCGCGGCAGCAGAACATTATCTTCAAGCGGTGCTATCTCTTGCTG	288
QY	167	ccatcgatccggggcttccctctcttgatgtgatcatcttcggttagctttcacccaaca	226
Db	289	CATGCGCCTCTGGAGCAGGAATTGGCGTTCAAACCCTCATCTTCGCCCACTTCTG	348
QY	227	agttcaacaattacaattcggcgccggaggacctcgaagcggttcaagccgagtgtgacc	286
Db	349	TCAATCACGATTTACCCACCGGAATCTCAACGCCGGCAGACTTTCGTGACATGCGCGC	408
QY	287	atttgg----tctgtggttgtctactctctttattcgggaagtgttgctctcagtaett	342
Db	409	AGTTGGCGTAATCTCTACTTTGTATCTTGGCATCTGCGGGGCTCTCTCTCTCTACACC	468
QY	343	tccacggctgcattaccatttcagcttatcgaaccactcgaactcttcgacgagtgttc	402
Db	469	TACAACACCCCTCTAACTACCGGGCCTACCGGCATCTCTCGCAATACTCGACACGCCAT	528
QY	403	cttgaatcaccttgggcgaaggtctggtcatcttcacaagcagacgaataatgaaccaatc	462
Db	529	CTCAAAGCGCGCTAGSCAAGAAGTGGCATATACGATTTTCGTATCGTAGCGGGGCTCCATC	588
QY	463	gccactcaatgctactaccaatggcaacogtatatacaaacaggtattgcgcagaaattggtc	522
Db	589	GCCGGCAGGCACTTTCGACGCAACTGTATCTACGAGCCGCCCTCTCGGATGAAGATCGT	648
QY	523	tttacccgtcaggaacatttoaatgttctttctgtcatttggctgccttcttgctgcgtcag	582
Db	649	CTTCTCTTCCAAGGGCTCGCAGCATCTGTGACGTTTCATATACGGGTTTTGTGTGCAAG	708
QY	583	tggagctgagctttaatcaccatgtccgttcactccctgccatttctcgtgcacgcgcatc	642
Db	709	TGGAACCTACACTCTGATCTGCATCTGCATCCCCGTAGSACCAGATCGGCACGAGGGGTG	768
QY	643	tgcatagcaattgatctccctcagagccaggaatcaccaagatatctatcagcgcgcgt	702
Db	769	GTAGCTGGGCTCGAGGCTGGGCACGAGACGAGGATCTTCAGATACATCGCAGCGCAAT	828
QY	703	gtccttcgagaagaagtctttaatacctccggacagtcacatgtcttctaagcccagaag	762
Db	829	TCGTTTGGCGAGGGATTTCTGGCGGCTGTGAAGGCTCTTCATNGCTTTTGGGATCGGGAT	888
QY	763	aaaatcgttcgaaaaatatgatgtcttttttgcagcaagcaccaagaagagggaagaaaa	822
Db	889	AGTCTGTGTAGGAAGTTTGAATATCTGCTGGAGCGCATTAAGTCTCGTAAGAAGATC	948
QY	823	tcgcacaaataatgggtctgttctcactagatcatttgcattttagcgtctactcgca	882
Db	949	TCGCGGCTGTGCTCTCTCTCGCGGAGTATACGATCATCTACTTTGGATNTGGG	1008
QY	883	cctggccttttggaaaagtttctcgaatgtatcagaatggcgaggttgcocagcttggcaaa	942
Db	1009	CTFGCGGTTTTGGCAGGGGATTCATATGTTTCGGCAGGGGGAGATTGGGACTCTGTGGGAT	1068
QY	943	gtcttactgttctctctcccgctacacttagcagccacgttccatctcaatgcttgcct	1002
Db	1069	ATCTTTACGGTTTTGTCTCTGCTGCTCATCTCGCTCAATCAACCTGACTTACTGCTGCGCG	1128







Db 2088 CCGCUACCAACCCUCCAUUCGCGGUGCAUAUGAACUCCAUAAAGGACCGCGACGAAUUAUGA 2147  
QY 2013 cgggactttgggtactcgtcctaataatgctactcctaataatgctactcctaataataaaaaa 2072  
Db 2148 GAACACAAAGACAAUGAUGGUGGCGGCCUUAUGUUCUUCUGGCGCGGAUUGUCCAGA 2207  
QY 2073 tctcactggtggtcttctgttgcataataaoggtttgatgatgctggtggtggtggtggtggt 2132  
Db 2208 ACUGAAUUGGCGCUAUCUGGUGGUGCAUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2267  
QY 2133 aggaacagccctttgttgcagattgctcactgtcttcctgagttgagttgagttgagttgagtt 2192  
Db 2268 CGGCCAAGCUAUCUUGAUGUGCGUGUGUGGAGGUGUUCACGUCUGGCGGAGAGCGUAU 2327  
QY 2193 acagaacggcgacacttttatattctgattgtttcttctgctgctgctgctgctgctgctgctg 2252  
Db 2328 GCUAGACAAAGGACUUCUUGGCGGUGCAUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2387  
QY 2253 aggtatttcacgattgttgcacatgcaacgttatttcaacagttgttcaacacatgctgcta 2312  
Db 2388 CUGUUAUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2447  
QY 2313 tcaagccgcaatgttccaaagagttactggtatcaagacatcgacatcctcctgacatcccgga 2372  
Db 2448 UCGAGCCGCUAUCUGGACGACAUUGUGCGACAGGAUAUCCAGUUCUUGACCGUGAAGA 2507  
QY 2373 gcaaatctctgctcactcagaaactgctcagctcactcactcactcactcactcactcactcact 2432  
Db 2508 GAACACUACGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2567  
QY 2433 gatataagcaaatctcttatattatctgttgcgtgtaaca--tcgtctcagcagtg 2490  
Db 2568 GAUGGCGUACAAACGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2627  
QY 2491 tctcactgactgactgactgactgactgactgactgactgactgactgactgactgactgactgact 2550  
Db 2628 CUGGCGCAUUGCAUUCUUGGAAACUAGGCGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2687  
QY 2551 cctgtttgtgactgactcactcagaaactggtggtggtggtggtggtggtggtggtggtggtggt 2610  
Db 2688 CCUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2747  
QY 2611 ggcacacttgcagaaagctgctggttgcagcagcagcagcagcagcagcagcagcagcagcagc 2670  
Db 2748 GAAGAAUAUUGGCGACGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2807  
QY 2671 ctcacttctgactcctcgaagccatgttctcgaacagctactcagcagcagcagcagcagcagcag 2730  
Db 2808 UUGG 2866  
QY 2731 cttgctaagatacctcgaagccttctgttggagcagcagcagcagcagcagcagcagcagcagcag 2790  
Db 2867 CUGUUCUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2925  
QY 2791 tcatgagatttttgctcctcctggtggtt--ttgtatgagcagcagcagcagcagcagcagcagc 2849  
Db 2926 DGAUUGG 2985  
QY 2850 ggtgagtagcacacaactcatttatatctcgttgcgtggtggtggtggtggtggtggtggtggt 2909  
Db 2986 GGGAGACUAGCAUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 3044  
QY 2910 agcag 2969  
Db 3045 GGGUG 3104  
QY 2970 ctatatcctcgtcgtcagcaattggaacgacacatcgcgaacagcagcagcagcagcagcagcag 3029  
Db 3105 CUACUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3164  
QY 3030 aaaaagccagtggtgagtcctcgtcagcctcagcagcagcagcagcagcagcagcagcagcag 3099  
Db 3165 UAAAGGCCCGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3224

QY 3090 acgtgattcggctcagttctccgcggtgttccatgacaatcgagccgaggaataattgt 3149  
Db 3225 ACGGCCAGACGCCUCCUUAUCCUAAAGGGGUGAUAUGAUAUAANAGGCGCAUAUUAU 3284  
QY 3150 agtttatgtggtggttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3209  
Db 3285 CGCUUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3344  
QY 3210 ctacgacccagcctcggggtggttcttctgacacagagaatattgcaagaatgtctgccc 3269  
Db 3345 CUACGAUCCAAACAAAGGAGCAUACAAUGCGGUCUCCACCCUCCACCCACCAUAAACCC 3404  
QY 3270 gctgtgtaccggtcgccatattgttcttcttcttcttcttcttcttcttcttcttcttcttct 3329  
Db 3405 CAUAUCCUACCGAAUAUUGGACUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3464  
QY 3330 cgttcgagagaatgtgacgttgcctcgaacgcaattatcaagaaga--gctttgtcaag 3388  
Db 3465 AAUACGGGAGCAUUAUCCUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3524  
QY 3389 gacgcttcccgcaagccaatgttctggttcttcttcttcttcttcttcttcttcttcttcttct 3448  
Db 3525 GUCGGCCUCCGCGCAGCUAAUUGCCUGGAGCUUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCU 3584  
QY 3449 aacgcttctgctcgaagcaggtgctgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3508  
Db 3585 CACGCGCGUGGUGUC--AGCGGGUCCCAACUCUCUGGGGGGCGAGCGGCAACGCAUUGCCA 3643  
QY 3509 tcgcaagacattgattgataatcaaatcgaactgttcttcttcttcttcttcttcttcttcttct 3568  
Db 3644 UUGCCCGCGGCUCAUCCCGAGAUCCAAAGAUCCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUC 3703  
QY 3569 tcgacacgcaatcgacgttctggttctggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 3622  
Db 3704 UGGAUACAGAGAGUGAGAGAUUCUGCAAGGUCUCUGAGGGGGGCGGCGAGCGGGG 3763  
QY 3623 gccaagcacaatagcagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3682  
Db 3764 ACCGGCUUAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG 3823  
QY 3683 ttgtgttgcacaggaatcgcaaacggaacggaacggaacggaacggaacggaacggaacggaac 3742  
Db 3824 GUGUAUUCUUGGAGGAGAAUUGGUGGAGUUGGUGGAGUUGGUGGAGUUGGUGGAGUUGUUAU 3883  
QY 3743 gaggagaatattcagagatgtgttggcacaatcttttagac 3783  
Db 3884 GGGGGUGUAGACGGAUGUGAGGCGGCGAGCGGCGGAGGAC 3924

## RESULT 4

US-09-352-552-3  
; Sequence 3' Application US/09352552  
; Patent No. 6050264  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrC of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:







Db 1981 TTGGTTGACCGGAGCGCTTATCCAAACTTGTGGAGGCTCAACGTATCAATGAACAG 2040  
QY 1862 -----aggacctggggctgacgaacaagaacatgagaagac 1902  
Db 2041 AAGGAAGTGCAGCGCTTGGAGGACCGCGCTGAGGATCTCACGAATACAGATATGCG 2100  
QY 1903 ctgcacgaagagcagcagagagctgctgctggaacacggcgaactgagcactcac 1962  
Db 2101 AAAATCAAACTGCTCAAGCGCATCATCCGATCTCGACGAAACCAACCAACCATGAC 2160  
QY 1963 accactgccacatcgaactgagagactggaagac-----ggaaggcgctgcg 2016  
Db 2161 CGCAGCGGACCCACACAGTCTGTTCACGCGGATCTTCTAAAGACCCCGGAACA 2220  
QY 2017 accttgggctact-----cgctctaaaaatgcataactcattgtcttcagaaaca 2067  
Db 2221 ACTCCGAATATCAATATGAGGCGCTGCTCAAAATTTGTCTTCAACCCGCTGAA 2280  
QY 2068 aaaaatctctactggtcttcttgggtcaacaataaagggtctctgatatgcgcggccaca 2127  
Db 2281 ATCCCGTACATGCTCATCGGCTGTCTTCTCAGCTGTAGCTGTGTGGTCCCAACCCAG 2340  
QY 2128 ttccaggaacaagcccttcttctcgagattgctcactgctctcgagttgagtgctcat 2187  
Db 2341 CAACGAGTGTATATGCTAAAGCCATCAGCACACTCTCGCTCCCGCAATATATAGC 2400  
QY 2188 gcgcgaaggaagcagagactttattctgattgtcttctgtcgtgctctaggaat 2247  
Db 2401 AAGCTTCACATGATGCGGATTTCTGCTCATGTGATGTTCTTCGTGTTGTATCATATCAG 2460  
QY 2248 ctgtaggatatctcagattggctggacatgaacgttatcttcacaagtgtgcacccat 2307  
Db 2461 TTTATACGCGAGTCAACCAATGCTGTGATTTGCGGTATGCTCCGAGACATATTCGT 2520  
QY 2308 cgtctatcaagcgcgaatgttccaacagtgactggatcaagacatcgaaactcctcgacatc 2367  
Db 2521 CGCGGAGAACGACTGCTTTCGAGGATACTCCGCTCAAGACATTTCTTCTTGACAAG 2580  
QY 2368 ccggagcaaatctctgctctcacatcgcaactgctcagctctaccacgcagtlgcaa 2427  
Db 2581 GAAGAGATAGCACCGGCGCTCTGACCTTTTCTCTCCACCGAGACGAGCATCTCC 2640  
QY 2428 gactgatatcagcaaatcttctcact--ttatatgttgcgttcgggtcaacatgctcagc 2485  
Db 2641 GGTGTAGCGGTGATCTAGGACGATCTTGATGACCTCCACACCTTAGGAGCGGT 2700  
QY 2486 agtgcctacacatagcctatgtagtgaactgggacctgggtgtgtgtgtgtgtgtaactt 2545  
Db 2701 ATCATTTATGCGCTGGCGATGGGTGGAATTTGGCTTAGTTGTATCTCGGTGTGCGG 2760  
QY 2546 ccacccctgttcttggctgacctcagaattcgtctagagacgaagctagaaagccgga 2605  
Db 2761 GTTCTCTGGCATGCGGTCTTACCGATTCTATATGCTAG--CCAGTTTCAATCAGCG 2817  
QY 2606 aactcggcaaatcttcagaaagtctggcttgcagcgaagcttaccgcgactcgcg 2665  
Db 2818 TCCAGCTTGTATAGGAGATCTGCAACCTTTGCTTGGAGGCTACATGCTATCCCG 2877  
QY 2666 accgtctctatttgaactcgaagccatggtctctcacaagctactcgaactgtgagc 2725  
Db 2878 ACAGTGGCTGATTAACCCGGGAAGGATGCTCGGAGATTTACCATGCGCCAGCT--TG 2935  
QY 2726 aagggtcttgaagatcatcacaagcttcttggctggcagatgttctgggttttcaactgt 2785  
Db 2936 AGCCACAAGGAGGACCGATCTAATCTCTCTTGTGAGTCTACCTCTGTATATGCTGCT 2995  
QY 2786 cacagctcagatggagtttggctattgcctt-ggattttgtattgagtcgataattgg 2844  
Db 2996 CGCAGGCACTGTTTCTCTGCTGCTCGCTGGGTTTGTAGGAGGACACTTCTTG 3055  
QY 2845 ctccaggtgagtaagacacaactcaatttatcatcttctggtgggcttctgttctgccc 2904  
Db 3056 GTACCCAGGATGATGACATTTTCGCTTCTTGTGTTCTTCCGAGATCTCTTGGT 3115

QY 2905 gtccaagcagcagcccccagttatttggcttactccacagagtttttaaccaaggtctcgctcgct 2964  
Db 3116 CTCAATCCCGGGGACCGCTCTTTTCCTTTGGACAGAG--CATGGGCAAGGCGAAGATGCG 3174  
QY 2965 ggaactatatactctctgctgctgagacattgaagcgacacatcccgaaacggagggagac 3024  
Db 3175 GCGCGCGAATTCGACGACTGTTTCGACCGAAAGCCACAAATGTGATCTCTTGAAGAG 3234  
QY 3025 aagaaagcagccagctgggtggtgctcgcctctcgacctcgagacattgaattcagttat 3084  
Db 3235 GCGGAGAAGCTCGAAACGGTGAAGGTGAATTCGAATTTAGGAACGCTGCATCTCAGATAC 3294  
QY 3085 cgtcaactgattcgtcgagttctccggggtttccatgacaactcgaacgaggaaca 3144  
Db 3295 CGACCCCGCCAGAACACCTGCTCTCGCGGCTTGGACCTGACCGTGAAGCCCTGGACAA 3354  
QY 3145 ttgtagcttatgtgggctgtctgctggaagtcgaactgaactgtgctgttgcgaa 3204  
Db 3355 TATGTGCGCTTGTGCGACCCAGCGGTTGTGGCAAGAGTACCACCATTTGCAATTCCTTGGAG 3414  
QY 3205 cgaattcagcccgacccctcgggccgaatttcatcttgcacagagaattattgcagaatg 3264  
Db 3415 CGCTTTTACGATGCGATTCGCGGTCCATCTTGTGTTGATGGGAAGGACATTAAGTAAACTA 3474  
QY 3265 tcgcccgtctgtaccgcgccatgtcttggctcaacaggaacccacacttaccaa 3324  
Db 3475 AATATCAACTCTACCCGAGCTTCTGTCTACCTGCTCAGCCAGCGCGACACTGTACCAG 3534  
QY 3325 ggctcgtctcgagagatgtgaactgtggcctcgaaagcgaattatcagaag-actttg 3383  
Db 3535 GGCACCATCAAGAAACATCTTACTTGTGTTTGTGCAAGATGACGTACCGAAGAAATTC 3594  
QY 3384 tcaaggaagccttcccgcaagggccaatgtcttggattttgtctatctcttaccagaaggc 3443  
Db 3595 TTGATTAAAGGTTGCAAGGACGCTTAATATCTACGACTTCATCATCTGCTCCCGAGGCG 3654  
QY 3444 ttgaaagccttgcgctcaacagagagtgagttctccgcgggcaacgacagcgagat 3503  
Db 3655 TTTAATACAGTTTGTGGCAGCAAG--GGAGGCAATGTTGCTGTGGCGCCAAAAGCAACGTGT 3713  
QY 3504 cgcctcgaagacattgattcgaataccaaagctgttctacttgactgacgagcagctc 3563  
Db 3714 GGCATATGCGCGAGCCCTTCTCGGATCCCAAAATCCTTCTTCGATGAAGCGAGTC 3773  
QY 3564 agcctcgaacgaaatcggaaactgtgttgaagctgcctcgatgagggcatccacag 3623  
Db 3774 AGCCCTCGACTCCGAGTCAGAAAAGGTCGTCCAGCGGCTTTGGATGCGCTGCCCGAGG 3833  
QY 3624 cgaacgacaatagcagtgccgacccgaacttccactattcggaaatgttgcattttt 3683  
Db 3834 CCGAACCAATTCGCGGTGTCACACCGACTCAGCAGATTCAAAAGGCGGACGTTATCTA 3893  
QY 3684 tgtgttgcacagggagaaatcgcggaacgggcaactcgcgggaactlacaacactgag 3743  
Db 3894 TGTTTTCGACCAAGCAAGATCTCGAAGCGGAACGACAGCACTGGTCCAGAAAAA 3953  
QY 3744 aggaagataattacagatg 3762  
Db 3954 GGCCCGTACTACGAGCTG 3972

RESULT 6  
US-08-996-545-3  
; Sequence 3, Application US/089996545  
; Patent No. 592898  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of  
; TITLE OF INVENTION: Aspergillus nidulans











APPLICANT: Skatrud, Paul L.  
 APPLICANT: de Waard, Maarten A.  
 APPLICANT: Peery, Robert B.  
 APPLICANT: Andrade, Alan C.  
 TITLE OF INVENTION: Multiple Drug Resistance Gene atrod of  
 TITLE OF INVENTION: Aspergillus nidulans  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Eli Lilly and Company  
 STREET: Lilly Corporate Center  
 CITY: Indianapolis  
 STATE: Indiana  
 COUNTRY: U.S.  
 ZIP: 46285  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/328.320  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/996.545  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Webster, Thomas D.  
 REGISTRATION NUMBER: 39,872  
 REFERENCE/DOCKET NUMBER: X-11766  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 317-276-3334  
 TELEFAX: 317-276-2763  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4002 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-09-328-320-3

[illegible]

Qy	622	attttctctggttcacccggcatctgtcactagcaattgatcgccgttcagagagcccgaggtatccacc	681
Db	784	CUGCUUUACACAGCGCGGUGUUCAGAUUUUAUUAUCAGUAACAGCAAAAAGUCGCUU	843
Qy	682	aggatctactcacgcgcgcgtgtctctcgcagaagaagctcttatccatcccgagcagctc	741
Db	844	GACAGCAACGGUGGACGCGGCACUGUUGUUGGAAAGGUAUCACAGCUCCACAGAAAUGCC	903
Qy	742	catgctttctcgcgccagagaataatggctcgaaaaatatgatctcttttgcgcaagca	801
Db	904	ACAGCGUUUGGCACCCACAGACAAGCUUGCGGAAGCAGUAUAGGUUACCUUACAGCAAGG	963
Qy	802	caccagaagaaggagaagaataatgccaaataatgsgtctgttctcaactgagtaettt	861
Db	964	GAGAAUUGGGACAAAGAACCAAGAUUUCAGUGGUUUUAUGAUUUGCGGCCCAUGUUUGGC	1023
Qy	862	tgcattacgtgtctatcgcactgyccttttgaaagggttttcgcgatgtatcagataatggc	921
Db	1024	CUUUAUGUACUGCAACUACGGUCUUGCUUUGGAUUGGUUUCUGUUCUGUGUAUGAUGGU	1083
Qy	922	gaggttgcgcagcttggcaaatcttcttactgtgttctcttccctacccttagcaccacg	981
Db	1084	GCAGU---CGAUGUGGGUGAUUUCACAGUUCUUAUGGCCAUUUGCAUUGCAUUGCUUC	1140
Qy	982	tcaatctcaatgctgcgcttccagcttcagctcgctttaccbaacgcgcgcatcttggctcc	1041
Db	1141	UCCUUUGGGACGUUAGUCCAAUUGCUAAGCAUUUAACACGUGUGGCCGCGCCGCA	1200
Qy	1042	gaattatcagtatcaattgccaacccacgcgcagctgcagccctctgcacccctcttgaaag	1101
Db	1201	AGAUAUUUGGACGACUAGCAUGCCAGUCCCAUUGAUUAGCAUUAUUCGAACGAMGGGAAG	1260
Qy	1102	cagccagagggctgcctaggctcaaatgtgatatccbaaacctggcatttgcctaccctctcc	1161
Db	1261	AGCUUGACCAUUGUUGAGGGCCCAUUGAUUAGCAUUAUUCGAUUAUUAUCCCAUCU	1320
Qy	1162	cgaccatctgcccgaagtacttcgagatttcaacttgcgaatctccagcttgccgacagcagc	1221
Db	1321	AGACCCGAGGUCACGCUAUGGAGGAUUAUUCUUGUUAUUGCCGCGGAGAAAACAACC	1380
Qy	1222	gcccctgctcgatgcagctagcgcgaagaagcacaatgctcgccgtctacttgaaecggtgg	1281
Db	1381	GCUUUAGUGCGCCCCUCUCGCGCUGGAAAAAGCGGUGGUGCGUUGUUGACGGAUUC	1440
Qy	1282	tatctgcacagctcggggagagataattctatggtgtggaactggggacacatacaatgtg	1341
Db	1441	UACAUUGCCUGUUCGCGUACGCUUUUGUAGUGGCCAUACACAUCAAGCAACCCUCAAUUC	1500
Qy	1342	aaatgggtgagaagccgcatctgcctcgttcaacaggaacctgtgtgttctcgtggcaaca	1401
Db	1501	CGUGCGCUUGGCCAACAGAUUCUUGUUUAGCCAGAGCGCUUUCUUUUUGGCACAGCAG	1560
Qy	1402	atctccagaacattgccaacgggtttctatgg---atggcaagagagatctgctctcgagaa	1458
Db	1561	AUUUAUAGAUAUUAAGGACGAGCUCAUCGCGCACAAAGUACGAGAAUAGAUUCCGAGGAU	1620
Qy	1459	aaacaaaatggagctgtgtcaaaaagctgtcaaaagcagcaatggcgactgttctattaat	1518
Db	1621	AAGGUCGCGGAACUACUCGAGAAAGCGGCAAAAUUGCGAAUUCUACUUAUUUAUACU	1680
Qy	1519	gagcttcgaacqgttatagactgaagtgtgcagcagccgcgcgccttgagtgaggt	1578
Db	1681	GCCUUUGCCUGAAGUUAAGACCAAUUGUUGGCGAGCGCUUUCUUCUUUACAGUUGGC	1740
Qy	1579	caacgacaaacgaattgaaatcgacacgaagtatcataatcgatcccaagaatcctgttactc	1638
Db	1741	CAGAAACAGCGCAUUGCAUUGCCGCGUGGUUAGUGCCCAACCAAAAUCCUUGCUCCUG	1800
Qy	1639	gatgaagctaccagcgcccttgaccgcgaagcgcgagaaagtgtatccagggagccttgaaac	1698
Db	1801	GAUAGAGUUAUUGCGCUUUGGACAAAUUCCGAAGCGGUGUUCAGACGACGACUUGGAG	1860
Qy	1699	cgaqltccaaagacgcactactttgttcattgcccaaaactagcactgtcaaaagt	1758

Db 1861 AGGGCAGUGAAGGCGGAACACAUUUGGUAUCGUCUACUCCUUUCCAGCAUCAAACG 1920  
QY 1759 gctggcaacatcgagtcatttcccagggggaataatctgtagcaaggaacacacacgaa 1818  
Db 1921 GGGCAACAUAUUGGUGUUGGUAUUGGUAUUGGUAUUGGUAUUGGUAUUGGUAUUGGUA 1980  
QY 1819 ttgatgaattcggtctgatttcacgcgactggtggtgac----- 1861  
Db 1981 UUGGUGACCGCGGAGCGGUUAUCGCAAAUUGGAGGCUCAACGUAUCAAUGAACAG 2040  
QY 1862 -----aggacctggggtgtagcaacaaagaacatgagagacc 1902  
Db 2041 AAGGAACGUGACCGUUGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100  
QY 1903 ctgagcaaaaggagcagcagagagtggtgtagcagcagcagcagcagcagcagcagcagc 1962  
Db 2101 AAAAUCAAAACUGGUGAAGCGCAUACUCCGAGUCGACGCGGAACCAACCAACCAUUGAC 2160  
QY 1963 accactgcacatctcaagctgagacacctggagaagc-----ggaagggtgcggtcg 2016  
Db 2161 CGCAGCGGACCCCAAGUGUGUUGCAGCGCGAUUUCUAAAGAACCCCGCAAAACA 2220  
QY 2017 acttgggtctact-----cgctctaaatgcactcaatcattgttctcagaaca 2067  
Db 2221 ACUCCGAUAUACUAUUGAGCGUGCUCAAAUUGUUGUUCUUAACCGCCUGAA 2280  
QY 2068 aaaaatctactagtgctctctgttgcacaaataaaggttctgatatagcgcgcaaca 2127  
Db 2281 AUCCCGUAUAGUCUACUGGUGUUGUUCUACUGUUGGUGGUGGUGGUGGUGGUGGUGGUG 2340  
QY 2128 ttccagagcaagcccttttgggttggagattgctcactgtctctggttgggtggtctat 2187  
Db 2341 CAAGCAGUGCUUAUUGCUAAGCAACGACACACUCUCGUCGCCAGAAUACAAUAGC 2400  
QY 2188 ggggcacaggaagcgggagacttttatctgtgattgttctgttctgtggtctaggaaat 2247  
Db 2401 AAGCUUCACAUAGUGGGAUUCUGGUAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2460  
QY 2248 ctatgagtagatttcaagattggtgagcagctgatttctacaaagttgtcaccat 2307  
Db 2461 UUUUAUCGCGAGUACCAUUGGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2520  
QY 2308 cgctatacagcgcaatgttccaaagagtagctgagatcaagacatgaactcctgacatc 2367  
Db 2521 CGCGCAGAGACGCGCUUCGCGAGCAUACUCCGUCAGAUUGUUGUUGUUGUUGUUGUUGU 2580  
QY 2368 ccggagcaaatctgtgtctcactgcacactgctcagctgtcactcaccacagcttgcaa 2427  
Db 2581 GAAGAGAAUAGCAGCGGCGUUGACCUUUCUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2640  
QY 2428 gattgatcagcaaatttctcat--ttatatgtgtgtggttcaacatgctctgagc 2485  
Db 2641 GGUUGUAGCGGUGAGUCUAGGACGACUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2700  
QY 2486 agtgcctcactagctatgataaactggcgtggtggtggtggtggtggtggtggtggtggt 2545  
Db 2701 AUCAUUUGCCUGGGAUUGGUGGUAUUGGUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 2760  
QY 2546 ccacccctgttctgtgtgctcactcagaattctgttagagaagaagctagaaagcgga 2605  
Db 2761 GUUUCUUGGGAUGCGGUUCUACCGAUUUAUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2817  
QY 2606 aactcggcaaatcttgagaaagtgtgggttgcagcagcagcagcttaccgcgagtcgg 2665  
Db 2818 UCCAAGCUUGUUGUAGGGAUUGCAAAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 2877  
QY 2666 accgtctcattgtgactctgagccatgttctccacagtagctcggacatgttgagc 2725  
Db 2878 ACAGUUGGUAUUAACCGGGAAGGGAUGUUGGAGAUUUAACCAUGCCAGCU--UG 2935  
QY 2726 aaggttctgataatcacaagcttctgttggagcagatgttctgttgggttctcattgt 2785

Db 2936 AGCCACACAGGACGAGCACAGUCUAAUUCUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 2995  
QY 2786 cacagcagatgagtttttggctattgccct-gggattttgttattgtagtcagtcgataattg 2844  
Db 2996 CGCGACCAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3055  
QY 2845 cttcaggtgagtagcagcaacactcaattttatatcatcttctgtggcgttctgttcttcc 2904  
Db 3056 GUCACACAGAGUAUACAUUUCGCUUUCUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3115  
QY 2905 gtccaagcagcagccagatattgagcttactccagagtttttaccagagctcgtcgcgt 2964  
Db 3116 CUCAAUCCGCGGGCACCUUUCUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3174  
QY 2965 ggaactatatactctggtcggaacttgaagccagaccatccgcgaacggaaggaagaaac 3024  
Db 3175 GCGCGGAUUCGCGAGACUUGCGGACGAAAGCCACAAUUGUAUACUUGGUGUGUAGAG 3234  
QY 3025 aagaaaaaggcccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3084  
Db 3235 GCGGAGAACGCGAAGCGUGGAAAGUAAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3294  
QY 3085 cgtcaacgtgattcggtcgagttctcgcggttctccatgacaaatcgagcaggaacaa 3144  
Db 3295 CGACCCGCGCAGACACGCUUGUUGCGCGCUUGGACCGUUGACCGUAGACCGUAGCAA 3354  
QY 3145 ttgttagcttagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3204  
Db 3355 UAGUUGGCUUGCGGACCGGCGGUGGAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 3414  
QY 3205 cgattctacgacccgacctggcggaatttctgttgcacagcaggaatattgcagaaatg 3264  
Db 3415 CGCUUUAUAGUUGUUGCGGCUUCCUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3474  
QY 3265 tcgcgctgttgcacggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3324  
Db 3475 AAUAUCAAUCCGACGAGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3534  
QY 3325 ggtcctcgttcggaatgtgagctgtggtggtggtggtggtggtggtggtggtggtggt 3383  
Db 3535 GGCACCAUCAAAGAAACUUAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3594  
QY 3384 teaaggagccttcgcaagcgaatggttgcattgttgcattgttgcattgttgcattgttgc 3443  
Db 3595 UUGUUAAGGCUUGCAAGGACGCUAUAUUCAGCUUACUUGUUGUUGUUGUUGUUGUUGU 3654  
QY 3444 ttgaaagccttgcggtcgaagcaggtgagtgcttctcgcggtggtggtggtggtggt 3503  
Db 3655 UUUUAUACAGUUGUUGGAGCAAG--GGAGGCAUUGUUGUUGUUGUUGUUGUUGUUGUUG 3713  
QY 3504 cgcatcgcagagcagcattgattcgaatccaaagctgttgcattgacgaagcagcgtc 3563  
Db 3714 GGCUAUUGCGCGCCUUCUUGGGAUCCAAAUCUUCUUCUUGUUGUUGUUGUUGUUGUUG 3773  
QY 3564 agcctcgaacgaatcgagcagctctggttcaagctgcctcgtcgtcgtcgtcgtcgtcgt 3623  
Db 3774 AGCCUUGACUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 3833  
QY 3624 cgaacgaacatagcagtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3683  
Db 3834 CGGAACCAUUGCGGUGGACACCGACACGACGACGACGACGACGACGACGACGACGACG 3893  
QY 3684 tgtgttgcacagcaggaatcgccgaacggcactcgcggaactacacagcagcagcag 3743  
Db 3894 UGUUUGCAACCAAGCAAGUUGUUGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3953  
QY 3744 aggaagatattcagatg 3762  
Db 3954 GCGCGGUUACUAGAGCUG 3972

Sequence 1, Application US/08612734B					
Patent No. 5914246					
GENERAL INFORMATION:					
APPLICANT: Peery, Robert B.					
APPLICANT: Skatrud, Paul L.					
TITLE OF INVENTION: Multiple Drug Resistance Gene of					
TITLE OF INVENTION: Aspergillus Fumigatus					
NUMBER OF SEQUENCES: 9					
CORRESPONDENCE ADDRESS:					
ADDRESSER: Eli Lilly and Company					
CITY: Lilly Corporate Center, DC1501					
STATE: Indiana					
COUNTRY: U.S.A.					
ZIP: 46285					
COMPUTER READABLE FORM:					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: Patent In Release #1.0, Version #1.30					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/08/612,734B					
FILING DATE: 08-MAR-1996					
CLASSIFICATION: 435					
ATTORNEY/AGENT INFORMATION:					
NAME: Craig, Anne I.					
REGISTRATION NUMBER: 32,976					
REFERENCE/DOCKET NUMBER: X-9681					
TELECOMMUNICATION INFORMATION:					
TELEPHONE: 617-354-9570					
TELEFAX: 617-354-4043					
INFORMATION FOR SEQ ID NO: 1:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 4047 base pairs					
TYPE: nucleic acid					
STRANDEDNESS: single					
TOPOLOGY: linear					
MOLECULE TYPE: other nucleic acid					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 1..4047					
US-08-612-734B-1					
Query Match 11.9%; Score 450.6; DB 2; Length 4047;					
Best Local Similarity 47.5%; Pred. No. 3.4e-138;					
Matches 1769; Conservative 0; Mismatches 1869; Indels 87; Gaps 11;					
QY	116	gatatacgcgaataatcacagactctcaatgtcatcgccgctcatctcgccatcgat	175		
DB	302	GCTATGCTCGGAAGAAGACATCTGTATACACTACGTCGGAGCGCAATCTGTGTATGCTG	361		
QY	176	cggggcttccccttccttgatgtcgatcatcttgcgttagcttcaacaaagttaaca	235		
DB	362	CTGGTGCCTCTCCCTTTGTTCAAAATCTCTCGGTTCACTTAGCTCAGGCTGCCAGG	421		
QY	236	attacaattcggcgacgggagtcctgaagcgtttcaaggccgatgtggatcattcgacc	295		
DB	422	GCATACTCTGGCTACTATGCCCTATCATGAGTCTATCAACAACACTAAGAAATGTGC	481		
QY	296	tgtggcttgtctaaccttttatgggaattgttcctcaagctacgtttccacggctgcca	355		
DB	482	TTTACTTTTGATCTCGGTATTCGGGATTTGTCACAGTCTATGTCACGACCCTGGGTT	541		
QY	356	ttaccaatttcagctatacagaacctcgaactcttcgaagagtggttccttgatgaacct	415		
DB	542	TCATTTATCTGGCGAACATCTCACAGAAAGATCCCTGAAAATATCTTTGAGGCTATCC	601		
QY	416	tgcggcaagaggtctggcatttcgacaaagcagcaatggagcaatcgccactcargtcaa	475		
DB	602	TGAGGCAGAAATATGGCTTACTTCGACAAGTTGGCGCGGTGAAGTTACACGCGTATCA	661		



Qy 3698 ggagaaatccgaacagcgactcagcgggaactacaacgactgagaggagatattacg 3757  
Db 3953 GCAGATCGTCGAAAGTGAACGACACGAGATTGATTCGAACAAAGGCCGGTATTACG 4012

Qy 3758 agatg 3762  
Db 4013 AGCTG 4017

RESULT 10  
US-08-612-734B-3  
: Sequence 3, Application US/08612734B  
: Patent No. 5914246  
: GENERAL INFORMATION:  
: APPLICANT: Peery, Robert B.  
: APPLICANT: Skatrud, Paul L.  
: APPLICANT: Tobin, Matthew B.  
: TITLE OF INVENTION: Multiple Drug Resistance Gene of  
: TITLE OF INVENTION: Aspergillus Fumigatus  
: NUMBER OF SEQUENCES: 9  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Eli Lilly and Company  
: STREET: Lilly Corporate Center, DC1501  
: CITY: Indianapolis  
: STATE: Indiana  
: COUNTRY: U.S.A.  
: ZIP: 46285  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent in Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/612,734B  
: FILING DATE: 08-MAR-1996  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Craig, Anne I.  
: REGISTRATION NUMBER: 32,976  
: REFERENCE/DOCKET NUMBER: X-9681  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 617-354-9570  
: TELEFAX: 617-354-4043  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 4800 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: other nucleic acid  
: US-08-612-734B-3

Query Match 9.7%; Score 367.4; DB 2; Length 4800;  
Best Local Similarity 46.7%; Pred. No. 1.4e-110;  
Matches 1730; Conservative 0; Mismatches 1826; Indels 147; Gaps 12;

Qy 199 tcatcatcttcggtagcttcacacaaagttcaacaaattacaattcgggagcggagat 258  
Db 690 TAGATTCTCTCGGTTCACCTAGCTTCAGCTTCAGGTCAGGTCATATCTCTGTTACTATGCCCC 749

Qy 259 cctgaagcgttcaagggcggatgtggtatcttcctcctgtgtgttcctacacctttatt 318  
Db 750 TATCATGAGTTCTATCACAACCTGACTAAGAAATGCTTTACTTTGTGTATCTCGTATT 809

Qy 319 ggggaatttgtctcacatgattttccacggctgccattaccatttcagctacagaac 378  
Db 810 GCCGAGTTTGTACAGCTATGTCAGCACCGTGGTTTCATTTATATCTGCGCAACATCTC 869

Qy 379 actcgaactctcagcaggtgttccttgaaatgcacaccttcggcgaagaggtctggcatttc 438  
Db 870 ACACAGAGATCCGGTGAATAATATCTTGAGGCTATCTGAGGCTATCTGAGGCAAGATATGCTTACTTC 929

Qy 439 gacaagcagagcaatggagcaatcgcgcactcargtcaactcaactcaggcaacccgtatcaaa 498  
Db 930 GACAAATGGGCGCGGTGAAGTATACACGCGTATACCTGCTGATACCAACCTGATTCAG 989

Qy 499 acaggtattccgagaaattggtctttaccgtgcagggcaacttttcaatgtctctttctgca 558  
Db 990 GACGCCATCTCTGAGAAAGTTGCTCTCACTTTGACCGCATTCGCCACATTTTAAACCGCA 1049

Qy 559 ttgtgtcgtctttggcgctcagtggaagctagcttttaatacaccatgctccgtctccct 618  
Db 1050 TTTATTTGCTGCGCTACGTCAGATATTCGAAGTTGGCTGATCTGTACCTCAACTATCCTC 1109

Qy 619 gcaattttccttggtccaccggcactctgcatagcaattgtagcgcgtcagggagggcagatc 678  
Db 1110 GCGCTGCTCATGTGTTATGGGAGGTGGGTGCGAGGTTTATTTGTGAAGTACAGCAAGAAATCT 1169

Qy 679 accaggtatctactcaacgcgcgcgtctcctcgcagagaagagttcttcaatccatccggaca 738  
Db 1170 AYTCAAAGTTATGCTGCTGCTGCAACTGTTCGGGAAGAAAGTCACTCAGCTCCATTCGGAAT 1229

Qy 739 gtccatgctttctacgccccagagaagaaatggtcgaaaaatagatgctttttcacagcaa 798  
Db 1230 GTFACCGCTTTTCGGCACTCAGGATAGCTTCGCCAACAATACAAACCACTCTGGCTGAG 1289

Qy 799 gccaccaagaagggaaghaaatacgccaaataatggsgtcttcttcttcaactagatc 858  
Db 1290 GCTGAAAAATGGGCGCTCAAAACAACAGGTCACTCCTTTGATGATGATTTGGTGTATGTC 1349

Qy 859 ttttgeatttaacgctgctatcgcaactggccttttggaaagggttttcgcgatgatacagaat 918  
Db 1350 GGTATCATGTTCTCGAACTATGCTCTCGGTTCTGATGGGATCTCGCTTCGTT---GTC 1406

Qy 919 gggaggttgcgagacttggcaagcttctactgttcttcttccgtcacctttagcagcc 978  
Db 1407 GGTAAAGAGTCAACGTGGGCCAAGTCTGACAGTTTGTGATGCTATCTGATCGGTTG 1466

Qy 979 acgtccatctcaatgcttgcgcttcaagttcagttcagttcgttttaccacgcgcgcattctggc 1038  
Db 1467 TTCAGTTTGGGCAACGTCGCCCAATGCTCAGGCTTTTACGAATGGTGTGCTGCGGCC 1526

Qy 1039 tcgcaattattcagtatcattgacaaacccagcagctcgcacccctctcgacctctctgga 1098  
Db 1527 GCGAAGATTTCAGAGCAGATTGACCCGACATCGCCACTCGACCCCTATTTCTGACGAAGG 1586

Qy 1099 aagcagcagagggcgtcctagtcgaattgagatccaaacccctggcacttgcctacccc 1158  
Db 1587 AAGTACTCGACCATTTTGAAGGAAATTCGAATTTCCGAATGTCAAAACATCTACCTT 1646

Qy 1159 tcccgacctctgcccaagtaacttcagatttcaacttgaacaaatcccgctggcgaagcgc 1218  
Db 1647 TCAGACCCCGAAGTTACAGTCATGGAAGATGCTCTCTTTATCGATGCCGCGGAAAGACT 1706

Qy 1219 acggccctcgtcgtgctacagtagcgcgaagcaaacacaaatggtcggtcttacttgaacgg 1278  
Db 1707 ACCGATTGTTGGGCCCATCTGCTCTGGAAGAGTACTGTGTGTCGCTTGGTGGAGGCG 1766

Qy 1279 tggatctgcacaggttcggggagatattactgtatggttgggaactgggacaatacaat 1338  
Db 1767 TTTTACCTTCCAGTAGAGGCCAGGTATTTGCTGGACGGCCATGATATCCAAACCCCTCAAC 1826

Qy 1339 gtgaatggctggaagccgcattcgcctcgttcaaacaggaacacctgtgttcttctgagc 1398  
Db 1827 CTCGGTTGGCTGGACAGCAGATCTCTCTTTGTCAGCCAGGAACCTGTTCTTTTTCAGCAC 1886

Qy 1399 acaactctccagacacattcccaacggtttcattgagtagagcaacgag---atctccctcgc 1455  
Db 1887 ACGATCTTTAGAAACATCGAACATGCTGCTGATTTGGCACCAAAATTCGACATGAGTCGAAG 1946

Qy 1456 gaaaaacaaatggagcttctgcaaaaaagcttgcaaaagcagcaatggcgacgtgttcaatt 1515  
Db 1947 GACAAAGATTAGAGGCTCTGTTGAGAAATGCGGCCAGAAATGTCATGATGATTTTATT 2006











Db	1651	TACGACCCCGCTAGTGGGTGTTGAAGACTTGACGGTGAAGATATACAGTTCGCGTTAAACCTC	1711
Qy	1342	aaatggctgagaagccgcatctgcgtctcaacaggaaacctgtgtgtttctgttggaaca	1401
Db	1711	AAATGGCTTCGCTCAACAGATGGCTGTGTATCTCAGGAACCCACTCTTTTCGGTACACCA	1770
Qy	1402	atcttcagaagacattgcacaacggtttcatgtagtagcaacagag---atctgcctcgggaa	1458
Db	1771	GTCCGCGGTAACTGTCGAACACGGCTCTCATGGCTCCAGATATACAGAAATGCCCTTCGCG	1830
Qy	1459	aaacaaatggagctgtgcaaaaagcttgcaaacgacgaatggcagcgtgttcaataat	1518
Db	1831	GAAAAATTTGAGCTTGTGAAGAAGGCTGTGTCGATGCCAACCCCATAAATTTTATCATG	1890
Qy	1519	gagcttcgaacggttatagactgaattggcagcagcgaacgagccttgagtgagagt	1578
Db	1891	AAACTTCCGCAAGGTTATGATACCATGTGTTGGTGAAGCGCGGTATGCTGTATCTGGTGGT	1950
Qy	1579	caacgaacaaatgtcaatcgcagcgaagtataatcggatccccaaagatccctgttactc	1638
Db	1951	CAGAACGACGAGTTGGGATGCTGTGCTATPGCTTTCGACACCTTAGAATTTGTGTGGT	2010
Qy	1639	gatgaagtaccagcgcccttgaccgaagcggcgagaaagtgtccagagagcccttgaaac	1698
Db	2011	GACGAAGCTACCTCTGCTCTTGACATCAGACGCAAGGTATTGTGCAAGATGCTTTGGAC	2070
Qy	1699	cgaagtctcaagaagccgcactcttggctcattgcccaaaactagccactgtcaaaagt	1758
Db	2071	AAGGCTCGCGAGCGAGGACTACTATCACCATTGCTCATAGGTTATCTACCAATTCGAGAT	2130
Qy	1759	gctggcaacatcgcagctcattctccaggggaaatcgtcagcaagcagcaacaacgaa	1818
Db	2131	GCGACACAAATTTACGTATGGGTGGTGGCGAAGTACTGGAGCAGGGTTCTCCRAACCT	2190
Qy	1819	ttgataagaattcg---gctgtcaattacgcgcacatggtgcgtgcacagacccctgggct	1875
Db	2191	CTTCTGCGCAACGAAACGGCTCTTACGCAACAACTTCTCAACAATCAAAAAGCTTGCTCAA	2250
Qy	1876	gacgaacaacagaacatcagagaagacctgcacgaaagcgcagcagagaagctgtcgtt	1935
Db	2251	GAAGCTGCTGTGAAGCCCTCCAGGTGGATGACGATATCGAGACCCCTCACCATGCCGTT	2310
Qy	1936	gaacgaccgycaccttgagcgccactcaacacccctgcacactctcaagctggagacccggag	1995
Db	2311	TTCAATTGGCGCTTCTTCGCTATGCAAGAAAAAGACAAACAGCTTCACTCGCGCAGTGACT	2370
Qy	1996	aagcgggaagtgccggtcgggaccttggctactcgtcctctaaatgcatacctaactcgt	2055
Db	2371	GGAAAGTCTTTGGCCAGTATCGCCATCGACGACATCCAACGCCAAGAGCGCGGAGGAATC	2430
Qy	2056	ttctacgaacaaaaatctctactggtgctcttctgttgtcaacataaacagctctcgata	2115
Db	2431	GCAGGTGAGGACAAAGATACCTCTCAAGCTTTGGTCTGTATGCCAGGCTGTTTAGGATGAAT	2490
Qy	2116	tgcgcggccacatt-----	2129
Db	2491	TCGCGCGATAAATTTATATACATCATAGCGTTTATCGCGCCCATCTGCGCAGGCATGGTT	2550
Qy	2130	--tccaggacaagccctttgtttctcgagatgctcaactgtcttcgagttgagtgagtgctcat	2187
Db	2551	TATCTCTTTGGCCATCTCTTTGGCAAGCCCTCTCCGACTTTTGAGATTTCAAGATCCT	2610
Qy	2188	gcgcacaggaacggcgagacttttatatt---cgaatgtctttctcgtgcgtcctagga	2244
Db	2611	GCGGAACCTTCACATGCTCTACGTCCGACTGCGCTGTGTACTTTTATCTACTGCTCTTGCG	2670
Qy	2245	aatctagtaggatatcttcacgatggcgtggacaatgcacaaogtatttccaaagttgcaccc	2304
Db	2671	GCTGATTTGTATCTTTCTTCCAAGATGCCGCGCTCTCTCTCGCCGCGGTTGGGATTTGAAC	2730
Qy	2305	catcgctatcaagccgcgaattgttccaaacgagctactggatacaagacatcgaaactcctgcac	2364
Db	2731	GGTGTCTTGAGAAAAAGTTTCTACCTGCTACTTTGGACATGACATCGAATGGTGTTCGAC	2790

Qy	2365	atccggagcaaaattcttcgggtgctctcaactgacactgctctctaccacgcagttg	2424
Db	2791	GAGGAACGAACCTACTCTGTGGCGTACATCAACCTTCCGATCAACCTCAGAAGTG	2850
Qy	2425	caagagttgatatacagcaaaatttcc--tcattatctgttgcgttcacacatgctctcg	2482
Db	2851	CAAGTCTGTGTGGGCTACTTTTGGGTACAGCTCGCTCCAGTCCCTGCCCACTCTCAATCGGT	2910
Qy	2483	agcagtgctcaccactagcctatgataagaaactgggctgggttgttgttggtagca	2542
Db	2911	GGTTGTATTATCGGCTTGTGCTACGGTCTCTTCTCTGCTGATCGTATCGCTTGT---	2967
Qy	2543	cttcacacctgcttggctgctactacacagaattcgtctacagacgaagctagaagcc	2602
Db	2968	ATCCCATCCCTTGTCTCGAGGTATATTGCTCTCAAGTCTGTTGTACTCAAGGACCA	3027
Qy	2603	gaaactcggcaaaactttgcagaaaagctcgggtcttcagagcaagacagctaccgcgac	2662
Db	3028	CGTATGAAGAAGCTCCATGCGGTAGTGCACATCTTGCCTAGTCAAGCTGCTGGAGCCGTT	3087
Qy	2663	cggacgctctcatctttgactctcgaagcgcatgtctccacagctactcggacatgtg	2722
Db	3088	AAGACAGTAGGCCAGCTCTTACCAGGGAGAGAGTGAAGGAGAATCTATTCTCAAGCTCTT	3147
Qy	2723	agcaaggtcttgcaagatcatcaaaagcttcttggtttgagcagatgttctgttttca	2782
Db	3148	A--NAGCTCCTATGAATTAATTTCCGTACTTCTATTAAATCCAGTGTCTCTTTCGGG	3205
Qy	2783	tgctcacgtcgatagagcttcttgctattgctccctggagatttgtatt--gcagtgcataat	2841
Db	3206	CGAGTCAGGTTGACTTCTCTGTATCATATTCCTGCTGCTGCTCTACATTTGGCGCATATTGGA	3265
Qy	2842	tggcttcaggtagtacyacacaaactcaatttatatactctcgttggcggtttgtttg	2901
Db	3266	TCATCGATTGCCAATATCTCCACATGCTAGTGTCTTACTACTGCTTGAACCTCCATGTGCTTTG	3325
Qy	2902	cgggtccaagcagcagccagtaattggcttactccacagttttaccagagctcggctcg	2961
Db	3326	-GCTCTATTCAAGCTGGAAATGCTTCACCTTCCTGCCGATGCTGCACAGGCCCAATATCC	3384
Qy	2962	gctgcgaactatctctgctgcggacattggaagcgcacattccgcga-----aaag	3015
Db	3385	TCGGCGCGCTCCATCTTCGAAGCATCGAATGAACCGGCTATCAATGCTGAATCTAAT	3444
Qy	3016	gaaggaacaagaaaaggcccgagtggtgagatgcctgctacactcagagacatgaa	3075
Db	3445	GAAGGGAGGTCCTTGATCACAGCACGTCGTGCTCATGTGGCGAATTGAGAGCGCTTCAC	3504
Qy	3076	ttaaggtatcgtcaacgtgatcggctcgagttcttcgcgggggtttccatgacaaacag	3135
Db	3505	TTTCAGGTATCTTACCAGCGCGGAGAGTACTGAGGAATTTGACTATTGATGTTCTCT	3564
Qy	3136	ccaggaacaatttgatctatgtgggcgctctctggctgtggcaagtcaacgttgatcgt	3195
Db	3565	GCGCGCACCTTATGTGGCACCTGCTGCTCTCCGTTGTGGAAAGTCTACCACTATTCAA	3624
Qy	3196	tigtgcgaagatctacgaccccgacctggcgcaatttcatttgcaacagagaatttt	3255
Db	3625	ATGCTGGAGAGTCTACGACCCCTTTTGGCTGGCGCTGTGACTTTGATGGCATCGATATT	3684
Qy	3256	gcagaaatgtcgcgcgcttgacgcgcgcacatgcttcttgctccacaggaacccaca	3315
Db	3685	AAGGAGCTCAACTTTGGGAGGTTACCTAGCCAGATCTCCCTGTTCTCAAGAACCTTACC	3744
Qy	3316	cttaccaaaggctccgtctcgagaaatgtagcctgttgccctcgaagccgaattatcagaa	3375
Db	3745	CTTTACGCGGGTAGAATTAGTGTTAATATCTTTTGGCGCGCAACAGCCCATAGAGAG	3804
Qy	3376	gagcttgttcgaagcgccttcccgc-----aaggcccaatgcttggatttctgctatc	3428
Db	3805	GTGACTCAAGGATGAGATTGCTGCTCGAAGGATGCCAATATTTCAGCACTTTATCTGT	3864

QY 3429 tctttaccagaggggtttgaaacgccttgcggctcaacgaggagtgacgttctccggcg 3488  
 Db 3865 TCCTCCAGATGGTTTGGACAC-CGAAGTGGGGGAAAGGCTCTCAGCTTTCTGGTGG 3923  
 QY 3489 gcaacgacgagctgcgcacgcgaagagcttgcgaatccaaagctgttgcact 3548  
 Db 3924 TCAAAAGCAACGTATCGCTATCGCCGCTCTCTCATTCGAAACCCCTAAAGCTTTGCTCCT 3983  
 QY 3549 tgcagagcagctcagccctgcacgcgaatcgaaagctgttgcactccctcga 3608  
 Db 3984 TGACGAGTACTGCTCGCAGCTCGACAGTCAGTCAGAAAAGCTGCTCAGAGGCACTTGA 4043  
 QY 3609 tgaggcatcacgagcgaagacacatagcagtggtgagcagcgaacttccactattcgaa 3668  
 Db 4044 TAAGGAGGCAAGGGTAGAACAACCTTGGCATTCGCTATTCGTCATTCACAA 4103  
 QY 3669 tgttgatgtattttgttggcgaacgggagaaatcgccgaacggggaactcacgcgga 3728  
 Db 4104 TTCTGACCGGATTTACTACTCTCTCTGAAAGCAGGCTAGCCGAGCATGGAACACATCAGGA 4163  
 QY 3729 actcagcagctgagagaagattacagagtg 3762  
 Db 4164 GCTTTGGCCAGAGAGGGTGGCTATTATGAGCTG 4197

RESULT 13  
 US-09-120-513-1  
 ; Sequence 1, Application US/09120513  
 ; Patent No. 6025160  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brun, Kimberly  
 ; APPLICANT: Chenery, Richard  
 ; APPLICANT: Ellens, Harma  
 ; APPLICANT: Field, John  
 ; APPLICANT: Yue, Lin  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
 ; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
 ; TITLE OF INVENTION: SCREENING METHODS THEREOF  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY:  
 ; ZIP: 19406  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; FILING DATE: 22-JUL-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: King, William T  
 ; REGISTRATION NUMBER: 30,954  
 ; REFERENCE/DOCKET NUMBER: GP50008  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-5015  
 ; TELEFAX: 610-270-5090  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4233 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Genomic DNA

Query Match 7.5%; Score 286; DB 3; Length 4233;  
 Best Local Similarity 44.9%; Pred. No. 1.2e-83;  
 Matches 1544; Conservative 2; Mismatches 1832; Indels 60; Gaps 10;  
 QY 360 cattcagctatcacgaccactcttcgacgagtggttcccttgatgcacttgcg 419  
 Db 430 CCTGGCAGCTGGGAGACAAATACACAAGATTAGCGAAGATTTTTCATGCATCATGAA 489  
 QY 420 gcaagagctctgcatttcgacaagcagagcaatgagcaatcgccactcargtccactac 479  
 Db 490 TCAGGAGATAGGCTGGTTGACCTGAATCAGCTGGGAGCTCAACACCCGGCTCACAGA 549  
 QY 480 caatggcaacggtatatacaaacagggtattgcccagagaaatggtcttccctgcagggact 539  
 Db 550 TGACGCTCTCCAAATTAATGACGGAATTTGGTGACAAATTTGAAATGTTTCTTTCAGTCCAT 609  
 QY 540 tcaatggtcttcttcgcatcttgcgtcttgcgtcttgcgtctcagtggaagctcagcttcaat 599  
 Db 610 AACGACATTTTCAGCCGGTTTATATAGGATTTATAGTGGTTGGAAGCTTAACCCCTGT 669  
 QY 600 caccatgcctgcctccctgcctatttctgcgtccacgcgcatctgcatacgacaattgatgc 659  
 Db 670 AATTTGGCGCTCAGCCCTCTTATTTGGGTTGTCTCATCTGCCATGTGGGCAAGGTACTGAC 729  
 QY 660 cgtcaggaggccaggtaccagggatctactcacgcgcgcgtctctcgcagagaagt 719  
 Db 730 TTCATTTACTTAATAAGGAACTCCAGGCTTTATGCGAAAGCTGGAGCAGTGGCCGAAGAT 789  
 QY 720 ctatcatccatccggagcagctccttcttccacccagagaaatggttcgaaaaata 779  
 Db 790 CTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAAAAGAGAACTTGAAGGTA 849  
 QY 780 tgaigtcttttgcagcagacacacagagaggaagagaaatcgccaaataatgsggt 839  
 Db 850 CAATAAATAATTTAGAACAGCTAAAGAGTTGCGCATTAAGAAAGCCATCAGCGCCAACAT 909  
 QY 840 ctgttctcactgagtgatcttgcatttcacgtctcactgcactgcacttcttggaagg 899  
 Db 910 TTCATAGGTTATTTGCTTACCTTGTGCTTATGCGCTTTATGCACTGGCATTTCTGGTATGG 969  
 QY 900 ttttcgcatgatcagaatggcgaggttgcgcagcttggcaagcttcttactgttgcct 959  
 Db 970 GACCTCCTTGGTCTCTCAATGAATATTC---TATTGGACAAAGTGTCTACCGCTTCTTT 1026  
 QY 960 ttcgctcacttagcagccactccatctcaatgccttcgccttcaggttcagtcgttta 1019  
 Db 1027 CTCTATTTTATTTGGGACCTTTTCAATTTAGGACATTTAGCCCAAAACATAGAAAGCTTTGC 1086  
 QY 1020 ccaacgcgcgcattcttcggtccgaaatattcagtatcatatgacaaacccacgcagctcga 1079  
 Db 1087 AAATGCAAGAGGGGCGAGCCTATGAAATCTTCAAGATAATTTGATATGAGCCCAAGCATTTGA 1146  
 QY 1080 cctctcgcgaccttctggaaagcagcagagggcgtgcctaggttcgaattgagatcccaaaa 1139  
 Db 1147 CAGCTTCTCAACCAAGGACACAAACCAGACAGTATATATGGAATTTGGAATTTAAATA 1206  
 QY 1140 cctggcatttgcctaccctccgaccattctcccaagtaacttcgagatttcaacttcac 1199  
 Db 1207 TGTTTACTTCAACTACCATACAGAGTGAAGTTAAGATCTTTGAAGGCCCTCAACCTGAA 1266  
 QY 1200 aattccagctggcaagcagcagccctcgtggtgcatacaggttagcggaagacacaaat 1259  
 Db 1267 GGTGAAGAGCGGCGCAGCGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAAACAGCAAC 1326  
 QY 1260 ggtcggcttacttgaacggtggtatctgcgcagttcggggaggaattacttgcgtggttt 1319  
 Db 1327 TGTCAGCTGCTGCAGAGGCTCTACGACCCCATAGAGGGCGGAGGTCTAGTATCGACGACA 1386  
 QY 1320 ggaactgggacaatacaatgtgaaatggtgcagagcgcagctcgttccaaagga 1379  
 Db 1387 GGACATCAGGACCATCAATGTGAGGTATCTGCGGGAATATCATTTGGGTGGTGAAGTCAGGA 1446  
 QY 1380 acctgtgtgttctcgtggcacaatcttccagaaacattgccaacggtttctcaggtgagca 1439

[illegible]

Db	2503	GGGCTCCAGGCTTGCCTGCTAGTTACCCAGAAATGTACAAACCTTTGGCGACAGGAATATCTTT	2506
Qy	2494	accactagcctatggatgataaactggccggctggctgtgttgggtgcacttccaccct	2503
Db	2563	ATCCTTTAGTCTATGGCTGCAGCTTACACTTTTACTTTAGTAATTT---ATACCACTTCAT	2619
Qy	2554	gcttttggctggctacactcagaattcgtctagagacgaagctagaagccggaactcggc	2613
Db	2620	TGCTCTGGGTGGAAATTTATTTGAAATGAAATGTTGTCTGGTCAAGCTTTGAAGGACGAAGA	2679
Qy	2614	aaactttgcagaaagtgcctgggcttgcgaagcagcagctaccgcgataccggagccgtctc	2673
Db	2680	AGAGCTAGAGATCTCTGGGAAGATCGCTACACAGCAATTTGAAACCTTTCCGCACTGTGTG	2739
Qy	2674	atctttgaactctcgaaagccatgtctccaaacagtaactcgaacatgttgagcaaggtctt	2733
Db	2740	CTCTTTGACTTGGGACGACGAAGTTTGAACCTATGTATGCCACAGACTTTGCAGATTAACATA	2799
Qy	2734	gctaagatcatccaaagcttttggtttggagcatgttttggtttcttcacttctcactgcg	2793
Db	2800	CAGAAATGCTTTTGAAGAAAGACACAGCTTTTG--GGATCACCTTGGCTCTCACCCAGAGCC	2857
Qy	2794	atggagtttttggctactccctcgggatttttattgcagtcgataatttggcttccagtg	2853
Db	2858	ATGATTTATTTTTCCTATGCTGCTGTTTTCCTCGGTTTCGGTGCCTACTTGGTGGCAGAGAA	2917
Qy	2854	agtacagacacaactcaattttatcatcttcgttgggcgttttggtttgcgcggtccaaaga	2913
Db	2918	CTCATGACGTTTGAAATGTTATGTTGGTATTTCTGCTGTGTCTTTGGTGGCATGGCA	2977
Qy	2914	gcagcccatgtattggcttactccagagattttaccaaagctcggctggctggaaatat	2973
Db	2978	GCAGGAAATACCAAGTTCATTTCGCTCTGACTACGCGAAGGCCAAGTCTCAGCATCCAC	3037
Qy	2974	atcctctggctgdcgacatgaaagccagaccatccgcgaaacggggagacaagaataaa	3033
Db	3038	ATCATGAGGATCATTCAGAAATTCGCCAGATTGACACTACAGCACGGAGGCTTGAAG	3097
Qy	3034	ggccagtggttgatggccctctgcacctcggagacattgaattcaggtatcgtcaacgt	3093
Db	3098	CCTAATTTGGTTTGAAGGAAATGTGAATTTAATGAGGATCATGTTCAACTATCCCAACCCGA	3157
Qy	3094	gattcggctcagctctccgcggggttccatgacaatcgagcagcaggaacttctgtagct	3153
Db	3158	CCCAACATCCACGTGCTTCAGGAGCTGAGCTTCGAGGTGAAGAGGGCAACACGCTTCGC	3217
Qy	3154	tatgtggggccttctggcttgaggcaagtcacagttgatcgttttgcggaacgattctac	3213
Db	3218	CTGGTGGGCAGCAGTGGCTTCGGGAAGAGTACGTGGTTCAGAGCTCTCGACGCGTTCTAC	3277
Qy	3214	gaccgacctcggggcgaatttcatattgcacacgagaatatggcagaaatgtgcgcgcgc	3273
Db	3278	AACCCCATGGCTGGAAACAGTGTCTTAGATGCCAAGAAATATAAACAACACTCAAGCTCCAG	3337
Qy	3274	ttgtacgcgggcataatgtcttttggctccaaacagggaacccacactttacccaaggtccggt	3333
Db	3338	TGGCTCGCGCCCACTGGGCAATTGTGCCAGAGAGCCCATCTCTGTTTGTACTGCAGGATC	3397
Qy	3334	cgcgagaatgtgacgttggccctcga-agccgaattatcagaagaagcttttcaaggacg	3392
Db	3398	ACCGAGAACATCGCTACGGAGACACAGCCGTGCTGTCTCATGAGGAGATCGTGAGG	3457
Qy	3393	ccctccgcgaagcgaatgctttggattttgtcatctctttaccagaagagcgtttgaaacg	3452
Db	3458	GCCGCGAGGGAGCCAAATCCACCNGTTTCATCGACTCCTGCTCTGAGAAATACACACAC-	3516
Qy	3453	ccttcggctcaacagggatgcagttctccggcggggcaacgacgacgaggtatcgccatcgc	3512
Db	3517	CAGAGTGGGAGACAAGAGGACATCAGCTGTCTGGATGAAGGACATCAGCTCTGGA	3576
Qy	3513	aagagcatgtattcgaatccaaagcgtttgctacttgacgaagcagcgtcagccctcga	3572
Db	3577	GCGGCCCTCGTTCAGACAGCTTCACATCTTACTTCTGGATGAAGGACATCAGCTCTGGA	3636



Db 1843 scaagaaatcatgaagagctcatgaagagaagggaacttacttcaaaactgtcatgac 1902  
Qy 1860 acaggacctcggggctgacgaacaacaagaacatgagaagacctgcacgaagaaggcagc 1919  
Db 1903 acgagactagaggaaaagaaatgaaccaggaaataatgtcttatgaatcccaagtgacac 1962  
Qy 1920 acgagaagctgtgggtg-----aaagccggcacttgagcgactcacac 1964  
Db 1963 tgggtcctcgtgacttacttcagaaaaatacaaatctctttaaagagatcaattcg 2022  
Qy 1965 cactgcacatctcaagctggagacctgagaagcgggaagtgccggctcgggactttggg 2024  
Db 2023 cagaagtatccacagaagaacagaccagagagaagacttaatttcgaagaagtgaga 2082  
Qy 2025 ctactgcctctaaaaatgcactaatatcatatgtttcacgaacaaaaatctctactggtg 2084  
Db 2083 tgaagtgtgcctatggtttcttttggcagatcctaaagctaaatattagtgagtgcc 2142  
Qy 2085 ctctcttg---ttgtcaacaataacggttctgtatgcgcggccacatttccagagacaagc 2141  
Db 2143 ctatttagttgggtgactttgtgctgtataaaatgggtgcatacaaacagtggttgc 2202  
Qy 2142 cctttgttttcagagattgctcacgtctt-----cgagttgagtggtcatgcggcaca 2195  
Db 2203 catagtggtttcaagatgtaggggttttttcaagagacagacacctgaacccaaca 2262  
Qy 2196 ggaacgggacagcttttatcttgatgttcttttgcgtgcgcgcacatttccagagacaagc 2255  
Db 2263 acggaattgtaactgtttctctctcttctgtgtatgggtggtatgattttctttgttac 2322  
Qy 2256 atatttcaagatggtgcgacatgaacgtttatttcaacgtgtgcacctgcctatca 2315  
Db 2323 gtacttcttcaaggcttcacatttggcaaaagctgagagatcctcaacaaagcactcgc 2382  
Qy 2316 agccgaatgttccaaagctgactggtatcaagacatcaaacctcgcacatcccgaggca 2375  
Db 2383 atacatggtctcaaatccatgctgcgacaggatataagtggtttgtagaccataaaa 2442  
Qy 2376 aattctggtgctcacatgcgaactgcagctctacccacgagtgacagagtgat 2435  
Db 2443 caccactggtcgtgactaccagctgctagtgacgcttcttaagttaaagggtctat 2502  
Qy 2436 atcagaataattctcatattatagttgtcg--gtcaacatgctctcgagcagtgctct 2493  
Db 2503 gggctcaggtgctgtagttaccagaaatgtagcaaaccttggcagaggaattatctt 2562  
Qy 2494 accactagctatggatgaaactggcctgggtggtgtgtgtgtgtggaacttccacct 2553  
Db 2563 atccttagtctatggtggcagcttacacttttactttagtaatt---ataccactcat 2619  
Qy 2554 gcttttgctgctacctcagaaattcgtctagagacgaagctagaagccggaactcggc 2613  
Db 2620 tgtcttgggctgaattatgaaatgaactgttctgtcgaagccttgaaggcaagaa 2679  
Qy 2614 aaacttgcgaagaagtgtcgggttcgaagcgaagcagttacccgcgacccgactctc 2673  
Db 2680 agaactagatctctgggaagatcgtctacagaagcaattgaaaacttccgcactgtgt 2739  
Qy 2674 atcttgaactcgaagggcactgttctccacagctactcggacatgttgagcaaggtctt 2733  
Db 2740 ctcttgaactagggacagaagtttgaaactatgtatgccagagcttgcagataccata 2799  
Qy 2734 gctaagatcatccaagcttttgggttggagcagatgttttgggttttcaacttgcacgtcg 2793  
Db 2800 cagaaatgctttgaagaagcacacgtctttg--ggatcacctcgtccttccaccagacc 2857  
Qy 2794 atggagtttttggctatgcccgtgggtatttattgtagctgcgataaattggcttcagtg 2853  
Db 2858 atgatttattttctatgctgctgtttcccggttcgggtgcctacttgggtggcagagaa 2917  
Qy 2854 agtacgacacaactcaatttatcatctctcgtgggggttttgggttcccggttccagca 2913  
Db 2918 ctoatgacgtttgaagaatgtatgttgggtattttctgtgtgtgttgggtggtcattgga 2977

Qy 2914 gacgccagttattggcttactccacagagttttaccagagctcgtcggtcgtcgcaactat 2973  
Db 2978 gcagggaataccagttcatctcgtccctcgaactacgcgaagcccaagctcagcatccac 3037  
Qy 2974 atctctggtcgtcggaacttgaagccgacctccgcgaacggagggagaacaagaaaaa 3033  
Db 3038 atcatcagagatcatgaaaaatcccccagattgacagctacagcacgggggttgaag 3097  
Qy 3034 gggccagtggttggtgctcgtcgtcagctcgtcagacattgaattcaggttctcgaact 3093  
Db 3098 cctaattggttagaagaaatgtgaaatttaattgaatggtcatgtttcaactatcccccga 3157  
Qy 3094 gattcggctcgtggttctccgcgggtttccatgacaatcagcagcagacaatttgaact 3153  
Db 3158 cccaacatcccaagtgcttcaggagctcagcttcaggtcgaaggggggcaacgcgtcgc 3217  
Qy 3154 tatgtggcgtcttctggtcgtggcaagtcaacttgccttctgtcgtcggaactctctac 3213  
Db 3218 ctggttgggcagcagtggtcgtgggaagatcacgtggtccagctgctcgcagcgtctac 3277  
Qy 3214 gaccgacctcggccgaatttctcgtcacagacagaaatttcagaaatgtcgcgcgcgc 3273  
Db 3278 aaccocatggctcggaaagtgcttctagatggcaagaataaaacaactcaacgtccag 3337  
Qy 3274 ttgtaccgggccaatgtcttgggtccacaggaacccaacttaccagaggtcttgaacgc 3333  
Db 3338 tggctccgcgccaactggcattgtgtcccagagccccatcctgtttgactcagcatc 3397  
Qy 3334 cgcgaatgtgactgtggcctcga--agccgaattatcagaagagcttctcgaagagc 3392  
Db 3398 accgagaacatcgtcactcagagacaacagcgtcgtcgtctcatgaggagatcgtgagg 3457  
Qy 3393 ctttcccgcaaggccaatgcttgggttttggatttttgcattcttaccagaggtcttgaacg 3452  
Db 3458 gcccacaggaggccaacatccacacagttcactcactcactgcctggaagaatacaaacac- 3516  
Qy 3453 ctttgcggtcacagaggtatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3512  
Db 3517 cagagtgggagacaagggactcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3576  
Qy 3513 aagaacttgcgaataccaaagctgttctactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3572  
Db 3577 ggcgctcgtcgaagagcctcactcttacttctggtgagacacatcagctcgtcgtcgtc 3636  
Qy 3573 cagcaatcggacatcgttgggttcaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3632  
Db 3637 tacgagagtgaaaaggctgtccagaaagcgtcgacaaagccagggggaagggcgacctg 3696  
Qy 3633 aatagcagtggtcgcgcgcacttccactatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3692  
Db 3697 catgtgactcgcgcgcgccttccactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3756  
Qy 3693 caacgggagaatcgcgaacacggcactcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3752  
Db 3757 gaacggcaggtcgaaggagcggccacccacagcagctcgtcgtcgtcgtcgtcgtcgtc 3816  
Qy 3753 ttacgagatgtgttggc 3770  
Db 3817 ttctcgtatggttcaggc 3834

## RESULT 15

US-08-181-471-2  
; Sequence 2, Application US/08181471  
; Patent No. 5641508  
; GENERAL INFORMATION:  
; APPLICANT: Li, Lingna  
; APPLICANT: Lishko, Valery K.  
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Thomas Fitting  
STREET: 12526 High Bluff Drive, Suite 300  
CITY: San Diego

COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/181,471  
APPLICATION NUMBER: 13-JAN-1994  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: ANT0039P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-792-3680

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4546 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:

NAME/KEY: CDS  
LOCATION: 425..4267  
3-181-471-2

Query Match	7.18;	Score 270.2;	DB 1;	Length 4646;
Best Local Similarity	44.9%;	Pred. No.2.le=78;		
Matches 1537;	Conservative	1;	Mismatches 1824;	Indels 63; Gaps 11;
QY	360	catttcagctatcagaacacactcgaactcttcagcagtggttctcttgaatgcacactctgcg	419	
Db	835	CCTGGCAGCTGAAGACAAAATACACAAAATTAGAAAACAGTTTTTTTCATGCTATAATGGC	894	
QY	420	gcaagaggcttgccatttcgacagcagcagcaatggagcaatgccactcargtcaactac	479	
Db	895	ACAGGAGATAGCTGGTTTGATGTGCAGATGTTGGGAGCTTAAACACCGACTTACAGA	954	
QY	480	caatggcacaacgtatacaaacaggtatgcgcgagaataatggctcttcacogtcagcgcaact	539	
Db	955	TGATGTCCTAAGATTAATGAAGTATATGTGTGACAAAATTGGAAATGTTCTTTTCAGTCAAT	1014	
QY	540	tcoaatgtctctttctgcatttggctggctcttggcgctccagtggaagcactagctttaat	599	
Db	1015	GGCAACATTTTTCACCTGGGTTTATAGTAGAATTTACAGTGGTGGGAAGCTAACCCCTGT	1074	
QY	600	caccatgtccgtccatcccctgcatttctcttggtccacgcgcactgcatagcgaatgatgc	659	
Db	1075	GATTTTGGCCATCAGTCTGCTTCTTGACCTCTCAGCTGCTGCTGGGCAAGATACTATC	1134	
QY	660	cgcctcaggagccaggatcaccaaggaatctactcacgcgcgcgtctcctgcgacaagaagt	719	
Db	1135	TTCATTTACTGATAAAGAACCTTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGT	1194	
QY	720	cttatcatcatcgcggacagtcctatgccttcacgccacagaagaataatggtcgaaaaata	779	
Db	1195	CTTGCACGCAATAGAACCTGTGATTTGCATTTGGAGGACAAAAGAAAGAACTTTGAAAGGTA	1254	
QY	780	tgatgtctcttttgcgacgaacacaccagaagggagaagaataatccgccaaa	839	



D	b	2308	GCAGACGACGAGGAATGAAGTTGAATTAAATAAATTCAGCTGATGAATCCAAAAGTGAAT	2367
Q	y	1915	gcagcacgagaagtctgtgtgaacgacggcacttgagcgcactcacacactgccaca	1974
D	b	2368	TGATGCTTTGGAAATGCTCTTCAATGATTCAAAGATCAGATCCTAATAAGAAAAAGATCAAC	2427
Q	y	1975	tctcaag---ctgaagacctggagaagcggaaaggctgcggtccggacittgggtact	2029
D	b	2428	TCGTAGAGAGTGCGGTGGATCAACAAGCCACACAGAAGCTTACTACCAAAGGGCTCT	2487
Q	y	2030	cgtcctaaaaatgatccc-----taatcatgttctacgaacaaaataatctcta	2078
D	b	2488	GGATGAAGATATACCTCCAGTTCTCTTTGGAGGATTAATGAAGCTAAATTTAACTGAATG	2547
Q	y	2079	ctggctctcttgttccaacaataaacogttcttatatgcgcggccacatttcagagaca	2138
D	b	2548	GCCTTATTATTGTGTGGTGATTTTGTGCCATTATAAATGGAGCGCTGCAACCAGCAT	2607
Q	y	2139	agccctttgttttcgagatgtctcaatgtctcttcgagttgagttggtcatgcggcacgga	2198
D	b	2608	TGCAATAATATTATTCAAGATTTATAGGGTTTTTACRAAGAATTGATGATCTCTGAACAAA	2667
Q	y	2199	acgggcagactttatatctgatgttcttctgctggtctcaggaaacttag-----t	2252
D	b	2668	ACGACAGATAGTAATTGTTTTCACATATTGTTCTAGCCCTTGGAAATTATTCTTTTAT	2727
Q	y	2253	agatatcttcagatgtggctggagactgaacggttatcttccaagttgtccccactcgcta	2312
D	b	2728	TACATTTTTTCCCTTAGGGTTTCACATTTTGGCAAGCTGGAGAGATCTCTACCAAGCGCT	2787
Q	y	2313	tcaagcgcgaatggttccaacsgtaactygtatcaagacatcgaaactctcgacatcccga	2372
D	b	2788	CCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGTGATGACCC	2847
Q	y	2373	gcaaatcttggtgctctcaactgcgaactgtcagctctcaaccaacgagttgcaagaagt	2432
D	b	2848	AACACCACTGGAGCATTCGACTACCAAGCTCGCCAATGATGCTGCTCAAGTTAAAGGGC	2907
Q	y	2433	gatatcagcaaatcttctc--atttatatcgttgtcggtcaacatgctctcgagcagtc	2490
D	b	2908	TATAGTTCCAGGCTTGCTGTAAATACCCAGAATATAGCAAACTTGGGACAGGAATAAT	2967
Q	y	2491	tctaacactagcctatgatgaaactgggcctgggtgtgtgtgttggtgcacttcacc	2550
D	b	2968	TATATCCTTCATCTATGTGTGGCAACTAACAC---TGTTACTCTTAGCAATGTGACCGAT	3024
Q	y	2551	ctgtcctttggctggttaoctcagaattcgtctagagacgaagctagaagcgggaaact	2610
D	b	3025	CATTGCAATAGCAGGAGTTGTTGAATGAAATGTGCTGGCAACGACTGAAAGATAA	3084
Q	y	2611	ggcaacttgcagaaaagtctgggcttgcaggcgaacgaactaccgcgatccggaccgt	2670
D	b	3085	GAAAGACATAGAAGGTCTGGGAAGATCGCTAGTGAACAAATAGAAACTTCCGAACCGT	3144
Q	y	2671	ctcatotttgactctcgaaggcatagtttccacaagctactcggacatgtttgagcaaagt	2730
D	b	3145	TGTTTCTTCTGACTCAGGAGCACGAAGTTTGAACATATGATGCTCAGAGTTT--CGAGTA	3202
Q	y	2731	cttgcataagatcatccaagcttttggtttggcagatgtttgttttctacttgtcaag	2790
D	b	3203	CCATACAGAAACTCTTTGAGGAACACACATCTTTGGAATTACATTTTCTTCACCCAG	3262
Q	y	2791	tcatgttgagtttttgggtatttgccctgggattttgttgcagtcgataatttggcttcag	2850
D	b	3263	GCAATGATGTATTTTTTCCCTATGCTCGATGCTTTCCGGTTTGGAGCTACTTGGTGGCACAT	3322
Q	y	2851	gtgagtacgacacaaactcaattttatatcatcttcgtgggcgtttttgttcgcggtccaa	2910
D	b	3323	AAACTCATGAGCTTGWAGGATGTTCTGTAGTATTTTCAGTGTGTGCTTTGTGTGCCATG	3382
Q	y	2911	gcagcagcccagttatttggcttactccacgagtttaccgaaggctcgtcggtgcgaac	2970
D	b	3383	GCCGTGGGCAAGCTGAGTTCAATTTGCTCTGACTATGCCAAGGCCAAATATCAGAGCC	3442

Search completed: April 27, 2002, 01:39:27  
Job time: 22456 sec

-----

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 00:42:14 ; Search time 706.01 Seconds

(without alignments)

9221.597 Million cell updates/sec

Title: US-09-882-694A-10

Perfect score: 3792

Sequence: 1 atggcagatgaatcgagaa.....aatcttagaccaagcatga 3792

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_032802:\*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3791.2	100.0	3792	22	AD02695
2	3454.8	91.1	3999	22	AD02695
3	723.6	19.1	3927	20	AX60999
4	722.2	19.0	3924	20	AX61000
5	487	12.8	4002	20	AX85823
6	450.6	11.9	4047	20	AX60201
7	436.6	11.5	8777	22	AXF83399
8	436.6	11.5	8777	22	AXF83400
9	367.4	9.7	4800	20	AX60202

10	362	9.5	3923	17	AAT35230	Aspergillus flavus
11	351.2	9.3	4224	19	AAV42347	cDNA encoding a mu
12	344.8	9.1	3133	22	AAF83398	P. chrysogenum ABC
13	319.8	8.4	7744	22	AAF83380	P. chrysogenum ABC
14	319.8	8.4	7744	22	AAF83381	P. chrysogenum ABC
15	309	8.1	3942	23	ABL14175	Drosophila melanog
16	306.2	8.1	4175	20	AAV69392	H. contortus PGP-A
17	300.2	7.9	4788	21	AAZ49335	Murine multidrug r
18	300.2	7.9	4788	21	AAZ49335	Mouse BCRP DNA rel
19	295.4	7.8	4369	21	AAZ52047	Rat multidrug resi
20	293.4	7.7	4425	21	AAZ52048	Rat multidrug resi
21	291.2	7.7	7476	22	AAF83393	P. chrysogenum ABC
22	291.2	7.7	7476	22	AAF83394	P. chrysogenum ABC
23	291	7.7	3924	21	AAZ94742	Human ATP binding
24	291	7.7	3924	21	AAZ94742	Human MDR-3 DNA.
25	290.4	7.7	4189	21	AAZ49334	Murine multidrug r
26	290.4	7.7	4189	21	AAZ49334	Mouse BCRP DNA rel
27	290.4	7.7	4313	14	AAQ38950	Mouse multidrug re
28	286	7.5	4233	21	AAZ90198	Rat mdrlb2 (multid
29	286	7.5	4233	22	AAZ27498	Rat mdrlb2 (multid
30	283	7.5	4279	22	AAQ03488	Dog P-glycoprotein
31	283	7.5	4279	22	AAQ03504	Dog P-glycoprotein
32	283	7.5	4279	22	AAQ03505	Dog P-glycoprotein
33	282.6	7.5	4317	22	AAQ03489	Dog P-glycoprotein
34	281.4	7.4	4279	22	AAQ03506	Dog P-glycoprotein
35	277.4	7.3	4186	22	AAZ86127	Cynomologous monke
36	277.4	7.3	4186	22	AAZ86128	Cynomologous monke
37	271.8	7.2	3860	21	AAZ49332	Human wild-type mu
38	271.8	7.2	3860	21	AAZ49332	Human BCRP DNA rel
39	271.8	7.2	3860	21	AAZ49332	Retroviral M4 ndr-
40	270.2	7.1	3860	21	AAZ49333	Human G185V mutat
41	270.2	7.1	3860	21	AAZ49333	Human BCRP DNA rel
42	270.2	7.1	4349	22	AAH57442	Human intestine ce
43	270.2	7.1	4646	15	AAQ72872	Human multidrug re
44	270.2	7.1	4646	21	AAZ94738	Human ATP binding
45	270.2	7.1	4669	8	AAZ70752	Sequence of human

#### ALIGNMENTS

RESULT 1

AD02695

ID AD02695 standard; DNA; 3792 BP.

XX

AC AD02695;

XX

DT 02-MAY-2001 (first entry)

XX

DE Exophiala spinifera p-glycoprotein coding sequence.

XX

KW Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;

KW aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;

KW detoxification; mycotoxin; animal feed; human feed; silage;

KW transgenic plant; transgenic animal; microbial spray; ds.

XX

OS Exophiala spinifera.

XX

PH Key Location/Qualifiers

FT CDS 1..3792

FT /tag= a

FT /product= "E. spinifera p-glycoprotein"

XX

PN WO200105980-A1.

XX

PD 25-JAN-2001.

XX

PF 14-JUL-1999; 99WO-US15824.

XX

PR 14-JUL-1999; 99WO-US15824.

XX

PA (PION-) PIONEER HI-BRED INT INC.

PA (CURA-) CURAGEN CORP.

XX Duvick JP, Maddox J, Gilliam J, Folkerts O, Crasta OR;  
XX WPI: 2001-147345/15.  
XX P-PSDB; AAY72637.  
XX Novel polynucleotides encoding Exophiala degradative or transport  
PT enzyme which is useful for detoxifying fumonisin or structurally  
PT related mycotoxin during processing of grain for human or animal food  
PT consumption  
XX  
XX Claim lb: Page 75-77; 90pp; English.  
XX  
CC The patent discloses novel polynucleotides encoding Exophiala spinifera  
CC fumonisin degradative or transport enzymes such as flavin monooxygenase,  
CC aldehyde dehydrogenase, permealase and p-glycoprotein. A fumonisin esterase  
CC enzyme with at least one fumonisin degradative enzyme is useful for  
CC detoxifying fumonisin or a structurally related mycotoxin present in  
CC harvested grain, on application to a plant or to harvested grain  
CC during processing, or to processed grain that is to be used as animal  
CC or human feed, or as a silage. Nucleotide fragments of the present  
CC invention are useful as probes and primers. They can be introduced  
CC into microorganisms that multiply on plants to deliver enzymes to  
CC potential target crops. The genes encoding the degrading enzymes are  
CC introduced via a vector into a microbial host and the transformed host  
CC is supplied to the environment, plants or animals for reducing the  
CC pathogenicity of a fungus producing fumonisin. The genes of the  
CC invention are fermented in a bacterial host and the resulting bacteria  
CC is processed and used as a microbial spray. The nucleotide sequences  
CC can be used alone or in combination to engineer microbes or other  
CC organisms to metabolise fumonisin and resist its toxic effects.  
CC The present DNA sequence is the p-glycoprotein coding sequence  
CC from Exophiala spinifera.  
XX  
XX Sequence 3792 BP; 939 A; 983 C; 928 G; 940 T; 2 other;

Query Match 100.0%; Score 3791.2; DB 22; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcagatgaatcgagaaacctcgaccacaaacgaatggcagtgctgctcacac 60  
DB 1 atggcagatgaatcgagaaacctcgaccacaaacgaatggcagtgctgctcacac 60  
QY 61 cctccccagaagaacggaagcagtgatttcagactatctacgaatcttcagatat 120  
DB 61 cctccccagaagaacggaagcagtgatttcagactatctacgaatcttcagatat 120  
QY 121 gccgacaaatagactgactctcaatgtctatcgctcatctcgccatcgatccggg 180  
DB 121 gccgacaaatagactgactctcaatgtctatcgctcatctcgccatcgatccggg 180  
QY 181 gcttccttcctctgatctgatctatctcggtgagcttcacaaagttcaacaattac 240  
DB 181 gcttccttcctctgatctgatctatctcggtgagcttcacaaagttcaacaattac 240  
QY 241 aattcgggacgagcgcctgaagcgttcaagccgatgtgatcatcttcgctgtgg 300  
DB 241 aattcgggacgagcgcctgaagcgttcaagccgatgtgatcatcttcgctgtgg 300  
QY 301 ttggtactctttatgggaagttgtgctcactgacgtttccacgggtgcattacc 360  
DB 301 ttggtactctttatgggaagttgtgctcactgacgtttccacgggtgcattacc 360  
QY 361 atttcagctatcagaccactcgactcttcgacagtggttcttgaatgcactggg 420  
DB 361 atttcagctatcagaccactcgactcttcgacagtggttcttgaatgcactggg 420  
QY 421 caagaggtctgcatcttcgacaaagcagagaatggagcaatgcactcargctacacc 480  
DB 421 caagaggtctgcatcttcgacaaagcagagaatggagcaatgcactcargctacacc 480

QY 481 aatggcaacggtatatacaaacaggtatgcccggagaaattggtttacacggtcagggcactt 540  
DB 481 aatggcaacggtatatacaaacaggtatgcccggagaaattggtttacacggtcagggcactt 540  
QY 541 tcaatgtttttctgcatttggctgcttggcgctccagtggaagctagctttaaacc 600  
DB 541 tcaatgtttttctgcatttggctgcttggcgctccagtggaagctagctttaaacc 600  
QY 601 accatgctcgtatccctccatttcttcggtcagccggaatctgcataagcaatgaagcc 660  
DB 601 accatgctcgtatccctccatttcttcggtcagccggaatctgcataagcaatgaagcc 660  
QY 661 gctcaggaggccaggatccacaggatctactcacgcgcgcgtctctcgcgagaagaagtc 720  
DB 661 gctcaggaggccaggatccacaggatctactcacgcgcgcgtctctcgcgagaagaagtc 720  
QY 721 ttatcatccatccggacagtcctatgctttctacgcccagaagaatggttcgaaaaatat 780  
DB 721 ttatcatccatccggacagtcctatgctttctacgcccagaagaatggttcgaaaaatat 780  
QY 781 gatgtctttttgcagcaacacacacaaagaggaagaagaatcgccaaatgaagagtc 840  
DB 781 gatgtctttttgcagcaacacacacaaagaggaagaagaatcgccaaatgaagagtc 840  
QY 841 ttgtctcaactgagtgacttttgcatttactcgtctgtatctgcactggccttttgaaaaggt 900  
DB 841 ttgtctcaactgagtgacttttgcatttactcgtctgtatctgcactggccttttgaaaaggt 900  
QY 901 ttctcgtatgtatcagaatggcaggttgcgcagcttggcagaagcttttactgtgtcctt 960  
DB 901 ttctcgtatgtatcagaatggcaggttgcgcagcttggcagaagcttttactgtgtcctt 960  
QY 961 tcggtcaacttagcagccagctccatctcaatgctgtgcgcttcagggttcaggttcattac 1020  
DB 961 tcggtcaacttagcagccagctccatctcaatgctgtgcgcttcagggttcaggttcattac 1020  
QY 1021 caacgcgcgactctcggtctccgaattattcagtatctattgacaaaccacacagctcgac 1080  
DB 1021 caacgcgcgactctcggtctccgaattattcagtatctattgacaaaccacacagctcgac 1080  
QY 1081 cctctgacccctcttgaaagcagccagagggctcctcaggttcaaatgagatccaaaac 1140  
DB 1081 cctctgacccctcttgaaagcagccagagggctcctcaggttcaaatgagatccaaaac 1140  
QY 1141 ctggcatttgcctacccctccgacactctgcccgaagcttcagatcttcaactgaaca 1200  
DB 1141 ctggcatttgcctacccctccgacactctgcccgaagcttcagatcttcaactgaaca 1200  
QY 1201 attccagctggcgaagcagcgccctcgtcgtgtgcatacaggtagcggcaaaagcacatg 1260  
DB 1201 attccagctggcgaagcagcgccctcgtcgtgtgcatacaggtagcggcaaaagcacatg 1260  
QY 1261 gtcggcttacttgaaacggttggtatctgccacagttccggggaggaatattacttgaggttg 1320  
DB 1261 gtcggcttacttgaaacggttggtatctgccacagttccggggaggaatattacttgaggttg 1320  
QY 1321 gaactgggacaaatacaatgtgaaatggctgagaagccgattccctcgttcaacagaaa 1380  
DB 1321 gaactgggacaaatacaatgtgaaatggctgagaagccgattccctcgttcaacagaaa 1380  
QY 1381 cctgtgtgtttctgctggcacaatcttccgaacatttgcacacgggtttcatgtagcaaa 1440  
DB 1381 cctgtgtgtttctgctggcacaatcttccgaacatttgcacacgggtttcatgtagcaaa 1440  
QY 1441 cgagatctgctcgcgcaaaaaaataatggagcttgcgaaaaagctcgcgaagccagcaat 1500  
DB 1441 cgagatctgctcgcgcaaaaaaataatggagcttgcgaaaaagctcgcgaagccagcaat 1500  
QY 1501 ggcagcgtgttcaatgaatgagcttcgaacaggttatgagactgaagtggcgagcagacc 1560  
DB 1501 ggcagcgtgttcaatgaatgagcttcgaacaggttatgagactgaagtggcgagcagacc 1560  
QY 1561 ggaagcttgagtgaggttcaacgacaaacgaattgtcaatcgacaggaagtatatcatatcgat 1620





QY 1077 cgaacct--ctcgacccttctgaaagcagccagagggtgccttaggtccaattgagatc. 1134  
Db 1100 cgaccctctcgaccctcttctgaaagcagccagagggtgccttaggtccaattgagatc 1159  
QY 1135 caaaacctggcatttgcctaccctctcccgaccatctgccaaagtacttcagatttcaac 1194  
Db 1160 caaaacctggcatttgcctaccctctcccgaccatctgccaaagtacttcagatttcaac 1219  
QY 1195 ttgacaattccagctggcagcagcggccctcgctcgctgcatcagctagcggcaaaagc 1254  
Db 1220 ttgacaattccagctggcagcagcggccctcgctcgctgcatcagctagcggcaaaagc 1279  
QY 1255 acaatggctggcttacttgaacgggtggtatctgctgccagctcgggagagatattacttgat 1314  
Db 1280 acaatggctggcttacttgaacgggtggtatctgctgccagctcgggagagatattacttgat 1339  
QY 1315 gggctggaaactgggacaaatacaatgtgaaatgctgagaagccgcatctcgctcgttcaa 1374  
Db 1340 gggctggaaactgggacaaatacaatgtgaaatgctgagaagccgcatctcgctcgttcaa 1399  
QY 1375 caggaaacctgtgttctgtggaacaattcttcagaaacattgccaacgggttctcatggat 1434  
Db 1400 caggaaacctgtgttctgtggaacaattcttcagaaacattgccaacgggttctcatggat 1459  
QY 1435 gagcaacgagatctgctcgcgaaacaaataaggagctgttgcaaaaagcttgcaaaagcc 1494  
Db 1460 gagcaacgagatctgctcgcgaaacaaataaggagctgttgcaaaaagcttgcaaaag-c 1518  
QY 1495 agcaatcgcgacgtgttcttaataatgagcttcgaaacggttatgagactgaagtggcgag 1554  
Db 1519 agcaatcgcgacgtgttcttaataatgagcttcgaaacggttatgagactgaagtggcgag 1578  
QY 1555 cgagcggagccttgagtgaggttcaacgacaaacgaatgtcaatcgcaacgaagtatcata 1614  
Db 1579 cgagcggagccttgagtgaggttcaac--aagccgaattgcaatcgcaacgaagtatcata 1637  
QY 1615 tcggatcccaagatcctgttactcgatgaagctaccagcgccttgaccgaaagggcgag 1674  
Db 1638 tcggatcccaagatcctgttactcgatgaagctaccagcgccttgaccgaaagggcgag 1697  
QY 1675 aaagtgttcagagccttgaaacgagtggtccaaagaccgactaatttggctatggcc 1734  
Db 1698 aaagtgttcagagccttgaaacgagtggtccaaagaccgactaatttggctatggcc 1757  
QY 1735 caaaaactagccactgtcatacgaactcaattagggcgaaattgggccccttagatgcat 1794  
Db 1758 caaaaactagccactgtcatacgaactcaattagggcgaaattgggccccttagatgcat 1817  
QY 1795 q-----tcgagcagcgacacaccccgaaattgatacgaaattcggtgtcattacgcgcac 1849  
Db 1818 gctcgagcgccgcagtgagcaattgatacgaaattcggtgtcattacgcgcac 1877  
QY 1850 tgggtcgctgcaagaccctcggggtgagcgaacaaacagaacacatgagaagaccctgcaag 1909  
Db 1878 tgggtcgctgcaagaccctcggggtgagcgaacaaacagaacacatgagaagaccctgcaag 1937  
QY 1910 aaaaaggcagcagagagctgctgggtgaacgaccggcacttgagcgcactcacaccactg 1969  
Db 1938 aaaaaggcagcagagagctgctgggtgaacgaccggcacttgagcgcactcacaccactg 1997  
QY 1970 ccacatctcaagctgagacacctggagaagcggaagtgccgtcgggactttgggtgaact 2029  
Db 1998 ccacatctcaagctgagacacctggagaagcggaagtgccgtcgggactttgggtgaact 2057  
QY 2030 cgctcctcaaatgcatctcaatgcttctacgaaacaaataatctctacttgggtgcttct 2089  
Db 2058 cgctcctcaaatgcatctcaatgcttctacgaaacaaataatctctacttgggtgcttct 2117  
QY 2090 tgttgtcaacaaataagcgttctgatatcgcgggccacatttccagggaacagcccttttgt 2149  
Db 2118 tgttgtcaacaaataagcgttctgatatcgcgggccacatttccagggaacagcccttttgt 2177  
QY 2150 tttcgagatgtctcaactgtcttcgagtttgagtggtgctacgagcaagcgagact 2209

Db 2178 tttcgagatgtctcaactgtcttcgagttgagttggtcatcgccaggaacgggcagact 2237  
QY 2210 tttatattctgatgtcttcttggctctaggaatacttagtagatatattcaagattg 2269  
Db 2238 tttatagttctgatgtcttcttggctctaggaatacttagtagatatattcaagattg 2297  
QY 2270 gctgagacatgcaacgttatttccaagtgtcccatcgctatcaagcgaatgttctc 2329  
Db 2298 gctgagacatgcaacgttatttccaagtgtcccatcgctatcgagcgaatgttctc 2357  
QY 2330 aacgagacttgatgataagagaatcgaaactctcgacatcccgagagcaaatttctgtgtc 2389  
Db 2358 aacgagacttgatgataagagaatcgaaactctcgacatcccgagagcaaatttctgtgtc 2417  
QY 2390 tcaatcgcaactgtcgactctaccacgagttgcaagttgataatcagcaaat--tt 2448  
Db 2418 tcaatcgcaactgtcgactctaccacgagttgcaagttgataatcagcaaatctct 2477  
QY 2449 ctcaatttatactgtgtcggtcaacatcgctcgacagatgctctaccactagccatgg 2508  
Db 2478 ctcaatttatactgtgtcggtcaacatcgctcgacagatgctctaccactagccatgg 2537  
QY 2509 atggaactggcctgggt 2568  
Db 2538 atggaactggcctgggt 2597  
QY 2569 cctcagaattcgtctagagcagagctagagcgcggaactcggcaactttgcgaaag 2628  
Db 2598 cctcagaattcgtctagagcagagctagagcgcggaactcggcaactttgcgaaag 2657  
QY 2629 tgcctgggctgtcagaagcagttaccgcatcccgagaccgtctcatctttgactctga 2688  
Db 2658 tgcctgggctgtcagaagcagttaccgcatcccgagaccgtctcatctttgactctga 2717  
QY 2689 aggcctatgtcttccaaacagtlactcggacatgttgagcaggttcttctaagatcatcaa 2748  
Db 2718 aggcctatgtcttccaaacagtlactcggacatgttgagcaggttcttctaagatcatcaa 2777  
QY 2749 agctttgtgtgtgacagatgtttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2808  
Db 2778 agc--ctgggtgtgagcagatgtttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2835  
QY 2809 attgcccctggattttgat---tgacgtcgataattggcttcagggtga--gtacgacaca 2864  
Db 2836 attgcccctggaaattttgt 2895  
QY 2865 actcaattttatactcttcgtggcggtttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2924  
Db 2896 actcaattttatactcttcgtggcggtttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2955  
QY 2925 tttggcttactccagagattttaccagaaggtcgggtcgggtcggaactatatactctgtgt 2984  
Db 2956 gccgaattactccacaggttaccagaaggtcgggtcgggtcggaactatatactctgtgtgt 3015  
QY 2985 ggggacattgaaagcgaccactcccgaaacggaggagagaacaaaaagggccagtgagg 3044  
Db 3016 ggggacattgaaagcgaccactcccgaaacggaggagagaacaaaaagggccagtgagg 3075  
QY 3045 tggatgcccctgacactcgagacattgaattcaggtatcgtcaacgttgatctggctctg 3104  
Db 3076 tggatgcccctgacactcgagacattgaattcaggtatcgtcaacgttgatctggctctg 3135  
QY 3105 agttctccgggggtttccatgacaaatcgagccagagcaaatttgtagcttatgtgggcgc 3164  
Db 3136 agttctccgggggtttccatgacaaatcgagccagagcaaatttgtagcttatgtgggcgc 3195  
QY 3165 tctcggtgtggcgaagccaactgt 3224  
Db 3196 tctcggtgtggcgaagccaactgt 3255  
QY 3225 ggggcgaatttcatcttgcacacgagaataatgtgcagaataatgtgcgcgctgtgtacgcgg 3284

Db 3256 gggcggaatttcatttgacacagagaaatttcagaaatgtcgccgcttgtaaccgcg 3315  
 QY 3285 ccataatgtcttgggtccaaacaggaaccccaactttaccaggctcgttcgcgagaaatgt 3344  
 Db 3316 ccataatgtcttgggtccaaacaggaaccccaactttaccaggctcgttcgcgagaaatgt 3375  
 QY 3345 gacgttgccctcgaagccgaattatcagaagagctttgtcaaggagcgccttcgcgcaag 3404  
 Db 3376 gacgttgccctcgaagccgaattatcagaagagctttgtcaaggagcgccttcgcgcaag 3435  
 QY 3405 gccaatgtcttggaatttgatctcttaccagaagagctttgaaacgccttcgcgctca 3464  
 Db 3436 gccaatgtcttggaatttgatctcttaccagaagagctttgaaacgccttcgcgctca 3495  
 QY 3465 acgaggatgcagttctccgcgggcaacgacgagcgatcgccatcgcaagagcattgat 3524  
 Db 3496 acgaggatgcagttctccgcgggcaacgacgagcgatcgccatcgcaagagcattgat 3555  
 QY 3525 tcgaatccaaagctgttgctacttgacgaagcgagctcagccctcgacacgaatcgga 3584  
 Db 3556 tcgaatccaaagctgttgctacttgacgaagcgagctcagccctcgacacgaatcgga 3615  
 QY 3585 acgtctgttcaagctccctcgatgaggcatccacgagccgacgaacgaatagcgtggc 3644  
 Db 3616 acgtctgttcaagctccctcgatgaggcatccacgagccgacgaacgaatagcgtggc 3675  
 QY 3645 gcacgactttccactattcggaattgtgatgttattttgtttgccaaacggagaaat 3704  
 Db 3676 gcacgactttccactattcggaattgtgatgttattttgtttgccaaacggagaaat 3735  
 QY 3705 cccgcaaacgagcactcacgcggaactacacgaactgagaggaatattacagagatgtg 3764  
 Db 3736 cccgcaaacgagcactcacgcggaactacacgaactgagaggaatattacagagatgtg 3795  
 QY 3765 ttggcacaacttttagacaaagatga 3792  
 Db 3796 ttggcacaacttttagacaaagatga 3823

## RESULT 3

AAx60999

ID AAx60999 standard; cDNA; 3927 BP.

XX AAx60999;

AC AAx60999;

DT 10-SEP-1999 (first entry)

XX A. nidulans atrC cDNA sequence.

XX Multiple drug resistance: atrC; Aspergillus nidulans; fungal disease;  
 gene disruption; ss.

XX Aspergillus nidulans.

OS Location/Qualifiers

FH Key 1..3297

FT CDS /\*tag= a

FT /\*product= "atrC"

XX W09932505-A1.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-US27499.

XX 23-DEC-1997; 97US-0996644.

XX (ELIL ) LILLY &amp; CO ELI.

XX (UYWA-) UNIV WAGENINGEN AGRIC.

XX Andrade AC, De Waard MA, Peery RB, Skatrud PL;

XX WPI; 1999-418904/35.

DR

XX P-PSDB; AAY21815.

PT New multiple drug resistance gene atrC

XX Claim 2: Page 21-30; 45pp; English.

PS

XX

The invention relates to a multiple drug resistance gene atrC of  
 CC Aspergillus nidulans. Host cells transformed with vectors comprising the  
 CC atrC gene are used for the recombinant expression of atrC polypeptide.  
 CC The atrC polynucleotide is used to produce nucleic acid molecules which  
 CC are useful in diagnostic assays for detection of atrC polynucleotides.  
 CC The atrC polypeptide may be used to identify agents which inhibit  
 CC multiple drug resistance. These agents may be useful for the treatment of  
 CC fungal diseases. New antifungal compounds may also be identified using  
 CC the gene disruption or gene replacement strains of Aspergillus nidulans.  
 CC The present sequence represents the cDNA encoding atrC.

XX Sequence 3927 BP; 856 A; 1040 C; 1132 G; 899 T; 0 other;

SQ

Query Match 19.1%; Score 723.6; DB 20; Length 3927;

Best Local Similarity 52.1%; Pred. No. 9e-217;

Matches 1929; Conservative 1; Mismatches 1745; Indels 29; Gaps 13;

QY 107 gaattctcagatatgccgacaataacgactggactctcaatgtcatcgcgtcatctgcg 166

Db 229 ggaattctacattcgccgagcagagaaactatctctcaggccgttgctactcgtctg 288

QY 167 ccactcgatccggggtctccctctctgtgtgtgatctcttcggtagcttcaccaca 226

Db 289 catgcctctggagcaggaattgcgttcaaaactatcttcggccagtctgcacgc 348

QY 227 agtcaacaattacaattcggcgagcgaggtctcgaagcttcaagccgagatgtgac 286

Db 349 tcataccagatttcaacaaacgggaattcgaagcggcgagacttctcgagaaatgccgc 408

QY 287 atttcg----tcctgtggtgtctacactctttattgggaagtgtgtctcactacgtt 342

Db 409 agtggcgatctctctactttgtatctctggcgatcgcgcggtctctctctacacc 468

QY 343 tcacgggtgcattacacattcagctatagcaacacactcgaactcttcgacgagtgttc 402

Db 469 tacaacacctcaacctacgcgcctaccgcctacgcctccgcaatattccgacgcctat 528

QY 403 ctggaatgcaccttcggcggaaggtctggcaatttcgacacagcagagcaatggagc 462

Db 529 ctcaagcggtcgctgagccaaagagtggaactactacgatttcggtagcgggggtcc 588

QY 463 gccactcargtcactaccatggcaaccggtatcacaaacagagttatgcgagaaaattggtc 522

Db 589 gccgcgaggcaacttcgaacggcaaaactgatccagggcgcgctcggtatgaatcggt 648

QY 523 tttaacgtcagcgcactttccaatgtctttctgcatttgcgtcttggtcgtctcag 582

Db 649 cttctctccagggtcgcagcatttcgacgcttctatcagcgtttgtggtgcaag 708

QY 583 tggaaagctagctttaaaccatgtctcctcatccctgcattttctcgtggttcacggc 642

Db 709 tggaaactcactctgatctgcatctgcattcccgtagccagatcggcagcgggggtg 768

QY 643 tgcatagcaattgatccgcctcagggccagagatcaccagagatctactcagcgcgcgt 702

Db 769 gtatgtcgggtcagggctgggcaagcagagagatcttgcaatcatatgcagcgcgaat 828

QY 703 gtctcgcagagaaggtcttatcatccatcccgagacagtcctatgcttctacgccagaag 762

Db 829 tegtgtccggagggttatctctggtgggtgtgaaggtcttcatgtcttctggggtcgggat 888

QY 763 aaaaatggcgaaataatgatgtcttttgcagcaagcacaccacgaaggaaggaagaaa 822

Db 889 agtctggtcaggaagttgatgaatctgtgtggagggcgcatgaaggtcgttcaagaatc 948

QY 823 tcgccaataatggsgtctgttctcaactgagtactttgtcatttactcgtcgtcatcgca 882





QY	3030	aaagggccagtggggtggatgcccgtgcagccctcgagacacattgaattcagggtatctca	3089
Db	3165	taaaggccctggatctggggcccgatgctatgacaaatgctgcctctcgtaacctct	3224
QY	3090	acgtgattggctcgagttctccgcggggtttccatgacaatcgagccaggaacaatttgt	3149
Db	3225	acggccagcgcctctatcctgaaggggtgaatctgaataaacaaggccaattcat	3284
QY	3150	agottatggggcctcttgctgtggcaagtcacgtgtgacgtttgtcggaaacgatt	3209
Db	3285	cgcttctgcgtccgtccgtggcgaaatccaccatgattccatgtctcgagcgtct	3344
QY	3210	ctacgaccgacctcgggcgcgaattctattgcacagagaatatgacgaatatgtcgcc	3269
Db	3345	ctacgatacaacaacaggagcatcacaatcgagcttccacctcaccgacataaaccc	3404
QY	3270	gccttgtaccgggccatgtctctttggctccaaacagaaacccacatttaccaggctc	3329
Db	3405	catatctaccgaatatcgtgcacctggtgcagaaagacacacacctttccaaaggac	3464
QY	3330	cgttcgcgagaattgacgcttggccctcgaagccgaattatccagaaga-gctttgtcaag	3388
Db	3465	aatacgggaacaatctcgcttggcgatgcagtgaagtcogtctgatgagcagattga	3524
QY	3389	gagcctctccgaagccaatcgcttggattgtgtctatctcttaccagaagcttga	3448
Db	3525	gtcgccctccgcgagctcaatgctctgggaccttctctctcatcttgcgcgaggga	3584
QY	3449	aacgcttgcgctcacgagggaatgcagttctccggcgggcaacagacagcgatcgcca	3508
Db	3585	caagcccgctggctc-aggcgggtcccaactcttgggggcagcggcgaacgattgcca	3643
QY	3509	tcccaagagcatgattcgaaatccaaagcttgcctacttgcaggaagcagcgtcagccc	3568
Db	3644	ttgccgcgcgtcatccagatccaaagatctactccttgacgaggtacgaagtgcgc	3703
QY	3569	tcgacacgaatcggaacctctggttccaagctgcctctcgatgaggcatcca-----cga	3622
Db	3704	tggatacagagatgaaagacatcgtcgaagaagctctctcgaggggcgccagggacgggg	3763
QY	3623	gccgaacgacaaatagcagttggcgaccgacttccactatttcggaattgattgtatttt	3682
Db	3764	acgggctacggtgtgctgtgcgcatgattaagcacgataaaggatgctaatgttatct	3823
QY	3683	ttgtgttgcacaagggaatatcgccgaacgggcacttccacgcggaactacaacgactga	3742
Db	3824	gtgtattctttggaggaagagattgvcggagatgggaacgcatcaagagcttaatagttaggg	3883
QY	3743	gaggaaagattaccagatcgtgtttggcacaatctttgaccaa	3786
Db	3884	gggggctgtataacgcaatgcttgagcgacgagcgtcttgactaa	3927

## RESULT 4

AA6100

ID AAX61000 standard: mRNA; 3924 BP.

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

AC AAX61000;

XX

DT 10-SEP-1999 (first entry)

XX

DE A. nidulans atrC mRNA

XX  
Muller + Co.

KW Multiple gene disruption

XX  
MY  
adnrtstn aileb

505

PF	22-DEC-1998;	98WO-US27499.	
XX			
PR	23-DEC-1997;	97US-0996644.	
XX			
PA	(ELIL ) LILLY & CO ELI.		
XX	(UWMA-) UNIV WAGENINGEN AGRIC.		
XX			
PI	Andrade AC, De Waard MA, Peery RB, Skatrud PL;		
XX			
DR	WPI; 1999-418904/35.		
DR	P-PSDB; AAY21815.		
XX			
PT	New multiple drug resistance gene atrC		
XX			
PS	Disclosure; Page 37-41; 45pp; English.		
XX			
CC	The invention relates to a multiple drug resistance gene atrC of		
CC	Aspergillus nidulans. Host cells transformed with vectors comprising the		
CC	atrC gene are used for the recombinant expression of atrC polypeptide.		
CC	The atrC polynucleotide is used to produce nucleic acid molecules which		
CC	are useful in diagnostic assays for detection of atrC polynucleotides.		
CC	The atrC polypeptide may be used to identify agents which inhibit		
CC	multiple drug resistance. These agents may be useful for the treatment of		
CC	fungal diseases. New antifungal compounds may also be identified using		
CC	the gene disruption or gene replacement strains of Aspergillus nidulans.		
CC	The present sequence represents the atrC mRNA sequence.		
XX			
SQ	Sequence 3924 BP: 854 A; 1040 C; 1132 G; 898 U; 0 other;		

Query Match	19.0%;	Score 722.2;	DB 20;	Length 3924;
Best Local Similarity	38.5%;	Pred. No. 2.5e-216;		
Matches 1426;	Conservative 502;	Mismatches 1744;	Indels 29;	Gaps 13;

Ov 107 qaatcttcagatatccgcacaaatacqaactqqaactctcaatqtcatcgcgcctcatctgcq 166

**F**

Db 229 ggaucuuacauccgccgcgcagggacagaaacuaaccuacggccguugcuaaccuugcug 288

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

**Qy**      ccacccggaaccggcgccctcccctcctgcagcgcgacccaccca  
T6 /                          :  
                      :

Db 289 caugcqcucucggaqcaqgaauuqcqcucuaaaccucacucucggaacccg 348

22

Qy 227 agttcaacaattacaattcgggacgggagtcctgaagcgttcaaggccgattgtggatc 286

109

b6 349 ucaucacccgauuucacaccacggaauucacacgcccgcaggaauucgugacaaugccgcgcg 408

Ov 287 atttccg----tccctatgaggttcgctctacccctctttattgqqaagttttatccctcacgtacqtt 342

[illegible]

Db 409 aguggcguaaccuacuuuguaaaccugggcaucgcgccuccuacacc 468

Qy 343 tccacggctgccattaccatttcagcctatacgaaccactcgaactctctcgaagtggtc 402

Dh 469 uacaaacaccilcciaaacccuaacagcagccuacccgcauuauccgcacacgccuuu 528

**QUESTION**

Qy 403 cttgaaatgcaccttgcggcagaggctctggcatctcgacaagcagagcaatggagcaatc 462

Db 529 cucaagcggcgcgcgaggaagggcauacuaacgaauucggguagcggggcuccauc 588



```
Db 2986 ggggagacagcaugauuaguuuugucgcauuccuacaguguuuuuugcggu-ca 3044
Qy 2910 agcagcagcagcatttcttactcagcagttttaccaggtcgtcgctgcgcaa 2969
Db 3045 ggcguacgagcagcuguuucagugucgaccagauuuaaaggggaaucuaucgacgaa 3104
Qy 2970 ctatctctcgtcgtcgacattgaagcgcacattccgcgaagcagaggaacaaaga 3029
Db 3105 cucaucgcuuguuugcaccagcucaacacagucgcgagacgcgagaaacaga 3164
Qy 3030 aaaagggccagtggtgtgctgcctcgcacccgcagcagcattgaattcagatcgca 3089
Db 3165 uaaagggccugaucugggcgccgaugcuaugcaauugcguucucgacccuccu 3224
Qy 3090 acgtgattcgcgtcagattctcgcgggttccatgacaatcgagccaggacaattgt 3149
Db 3225 acgcccagaccccccuaucuccuaaggggugaucugaauaacaagggccaaucuu 3284
Qy 3150 agcttatggcgcttctgctgtgcaagtcacagctgctgttctcgcgaacgatt 3209
Db 3285 cguuuucgucgucucccgucgagcgaauccaccacgauguccaugcagcgcuu 3344
Qy 3210 ctacgaccgacctcgccgcaatttcatttcacacagagaatttgcaaaatcgcc 3269
Db 3345 cuacgacccaacaacagggagcaucauacagcguuuccaccuccacgacauaacc 3404
Qy 3270 gcgtctgtaccgcgcacatgtctgttggccaaaggaaacacactttaccaggctc 3329
Db 3405 cauaucuccgaaauuugggcacugugcagcagagagccaaacccuuuuccaaggag 3464
Qy 3330 cgttcgcgagaaatgacgttgcctcgaagcgcgaattatcagaaga-gctttgtcaag 3388
Db 3465 aaucagggaacaacucucuuugcggaugcagugagucgucugaugacagaaua 3524
Qy 3389 gacgcttcgcgaagccaatgcttggatttctgtctatcttaccagagggcttga 3448
Db 3525 gucggccucgcgcagcauagccugggacuuuugucuccuauugcgcaggggaucua 3584
Qy 3449 aacgcttgcgtcgaagaggatgcagttctccggcgggcgaacgacagcgatcgcca 3508
Db 3585 caegcccgucguc-aggcggucccaacucucugggcgagcggaacauugcca 3643
Qy 3509 tcgcaagagcattgattcgaaatccaaagctgtgtacttcagcagcagcagcagccc 3568
Db 3644 ugfcgcgcgcuaucgcgagacuaacaaagaaucuaucuccuagcagaggcuagugccc 3703
Qy 3569 tcgacagcaatcggaacgtctggttcaagctgcctcctcagataggatcca----cga 3622
Db 3704 uggauacagagagugagaagucgucgagaaggcucucgagggggcgccaggagcgggg 3763
Qy 3623 gccgaacgacaatagcagtggcgcacgcactttccactctcgaatgtgtatgtatt 3682
Db 3764 accggcuacgguugucgcaucgaaucgaauaagcagcauuaagggaugcuauguaucuu 3823
Qy 3683 ttgtgttgcgaagagaaatcgccgaacgggacactacgcgcgaactcgaacgactga 3742
Db 3824 guuaucuuuugggagaaagucgagagagggagcgaucacgaagauuaauuaggg 3883
Qy 3743 gaggaagatatcagatgtgttggcacaactctttagac 3783
Db 3884 gggggcuuagacgcuugugagggcgagccuugac 3924
```

## RESULT 5

AAx85823

ID AAx85823 standard; cDNA; 4002 BP.

XX

AC AAx85823;

XX

DT 10-SEP-1999 (first entry)

XX

DE cDNA encoding multiple drug resistance gene atrD.

XX Multiple drug resistance; atrD; fungal disease; antifungal; ss.  
KW Aspergillus nidulans.

OS Location/Qualifiers  
FH 1..4002

FT CDS  
FT /\*tag= a

XX /note= "no termination codon given"

PN WO9932608-A1.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-US27300.

XX 23-DEC-1997; 97US-0996545.

XX (ELIL ) LILLY & CO ELI.

XX (UYWA-) UNIV WAGENINGEN AGRIC.

XX Andrade AC, De Waard MA, Peery RB, Skatrud PL;

XX WPI; 1999-418917/35.

XX P-PSDB; AAY02630.

XX New multiple drug resistance gene atrD

XX Claim 2; Page 21-27; 35pp; English.

XX The present sequence encodes a multiple drug resistance protein, designated atrD, which is isolated from Aspergillus nidulans. The atrD polynucleotide is used to produce nucleic acid molecules which are useful in diagnostic assays for the detection of atrD polynucleotides. The atrD polypeptide may be used to identify agents which inhibit multiple drug resistance. These agents may be useful for the treatment of fungal diseases. New antifungal compounds may also be identified using the gene disruption or gene replacement strains of Aspergillus nidulans.

XX Sequence 4002 BP; 974 A; 1034 C; 1007 G; 987 T; 0 other;

XX Query Match 12.8%; Score 487; DB 20; Length 4002;

XX Best Local Similarity 48.6%; Pred. No. 4.2e-142;

XX Matches 1731; Conservative 0; Mismatches 1760; Indels 68; Gaps 12;

Qy 262 gaagcgttcaagccgagtgatcatcttcctcgtgtgttcgtctactctttattggg 321

Db 424 gacgagttctatgatgaattgaccaagaacgtactgtactctgtatacctcggatcgc 483

Qy 322 aagtttgtctcagctacgtatttccacggctgcattaccattcagctatcgaacccact 381

Db 484 gatttgtcacgtctatgttagtactgttggtctcattatcgcgagagaacagccag 543

Qy 382 cgaactcttcgacgagtgctcttgaatgcaccttcggtgcggaaggggtctggtcattcgac 441

Db 544 cagaagatccgcgagattacattgagtcctcctcgtcgcgacacattggctatttgat 603

Qy 442 aagcagagcaatgaggaacacccactcactcactacacattggcaacccgatatcaaca 501

Db 604 aaactcgggtgcggggaagtgaccacccgataaagcgcgatacaaaccttctcaggat 663

Qy 502 ggtattgcgcgagaaaattggtcttaccctgcaggcacttcaatgttcttcttgcatttt 561

Db 664 ggcatttcgagaaggtcgtgtctcattcttgcctcgtcgcgacattgtgacagcttc 723

Qy 562 gtggtcgtcttggtcgtcagtggaagcagctgcttaatacaccatgtccgttcacccctgc 621

Db 724 attatgcctacgtcaaatcactggaagttggtcttaattgacgtcaacaattgtgccc 783

Qy 622 atttctcgtgcagcggaatcgtcgaatgacaaattgattgcgtcaggtcaggtcaggtaccc 681

Db 784 ctggttccaccatggcggtgtgttctcagtttatacacaagtaacagcaaaaagtcgctt 843  
QY 682 agatatactacacgcccgtctctcgcagaagaagcttatactatccatccgcagacgtc 741  
Db 844 gacagctacggtcagcgccgactgttcgggaagagtcacacgctccatcagaaaagcc 903  
QY 742 catgctttacgcccagagaanaatggtcgaaaatatgatgttttttttcagcaagca 801  
Db 904 acagcgttttggcaccacaagacagcttgcgaagcagtagtggtccactagacgaagct 963  
QY 802 caccagaaggggaagaaaatcgccaaaatgsgtctgttctcactaacatgagtaacttt 861  
Db 964 gagaaatgggaacaagaaccagatgtcatgggttttcagattggcgcccatgttggc 1023  
QY 862 tgatattaagctctatcgcactggcccttttggaaaagtttttcgatgatcaagaatggc 921  
Db 1024 ctatgtactcgaactacaggtctgtgtgttctggatgggttctctgttctctggttagtgg 1083  
QY 922 gagggtgcgacgttggcaaaagcttttactgtgttcctttccgtaaccttagcagccacg 981  
Db 1084 gcagt---cgatgggtgatattctcacagttctcagttctcagtcggaatctgtgatcgttc 1140  
QY 982 tcatctcaaatgcttgcgcttcaggttcagctcgttttaccacgcgcgcacatcttcggctcc 1041  
Db 1141 tcttgggaagcgttagtccaaatgctcaagcatttacaacgcgtgtggcgcgccgcga 1200  
QY 1042 gaattattcagtatcatgacaaaaccacgcagctcgcacctctcgaaccttctggaaaag 1101  
Db 1201 aagatatggaaagcatcgatcccgatcccatattagatccatatattcgaagcgaaggga 1260  
QY 1102 cagccagagggctgcctaagtcnaattgagatccaaaaccttggcatttgcctacacctcc 1161  
Db 1261 acgtcgaccattttagggccacatttgatttaccgatactatgccaagcatttaccatct 1320  
QY 1162 cgaccatctgccaaagtaactcagatttcaacttgacaattccagctggcgaagcagcag 1221  
Db 1321 agaccgaggtcacctcagtgaggtgttctctgtcaatgcgcgtggaaaacacc 1380  
QY 1222 gcctcggtcgtcatcaggtagcgaagaaagcaaatggtcgcgttacttgaagcgtgg 1281  
Db 1381 gctttagggccctctggtctggaagaaagcacggcggcgttgggtgagcgttc 1440  
QY 1282 tatctgccagttcggggaggtatattactgtatgggttggaaactggggacatacaatgtg 1341  
Db 1441 tacatgcctgttcgcggtcaggttttgggtggccatgacatacgaagcactcaatctc 1500  
QY 1342 aaatgggtgagaagcgcgcatcgcctcgttcacagacacctgtgttcttctgtggcaca 1401  
Db 1501 cgtggttcgcacaagatcttgggttagcgaagagcgtgttcttcttggcagcagc 1560  
QY 1402 atcttcagaaacttgccaaagtttctatgg---atggaaacagagatcgtcctcgcaa 1458  
Db 1561 attataagaataattagcagcgtctcctcgcgaagagtagcgaagatgaatccgagat 1620  
QY 1459 aaacaaatgagctgttgcaaaaagcttgcaaacgcagcaaatggcgagcgtgttctatcaat 1518  
Db 1621 aaggtccggaaactcagcagacgcggcaaaaatggcgaaatgctcagacttttact 1680  
QY 1519 gagcttcgaacgcttatagactgaagttggcgagcagcgcgcgcgtttagtggaggt 1578  
Db 1681 gccttgcctgaaggttatgagaccatgttggcgagcgtgttcttcttccaggtggc 1740  
QY 1579 caacgacaacgaattgcaatcgaagatatacattcgcgtcccaagatcctgttactc 1638  
Db 1741 cagaaacagcgcattgcaatcccggtgcgtttagtgaccccaaaaatcctgctcctg 1800  
QY 1639 gatgaagctaccagcccttaccgcgaagggcggaagaaagtggtccaggaagccttgaac 1698  
Db 1801 gatgaagctacttcggccttgacacaaaatccgaagcgtgttcttcaagcagctttggag 1860  
QY 1699 cgaggttccaaagacgcgactacttggctcattgcccacaaactgacagcctgtcaaaagt 1758  
Db 1861 agggcagctgaagcgcaactactattgtgatcgtcgtcgtcgtcgttcttcacacgatcaaacg 1920

QY 1759 gctggcaacatcgagtcatttccacaggggaaaatcgtcgagcaagcagcacacacagca 1818  
Db 1921 ggcacaacatgtgtgttctgtgtaattgctgaacaaaggaactcagatgaa 1980  
QY 1819 ttgatogaaattcggtctcatatcagcgcgcacgtgtgtgtgcac----- 1861  
Db 1981 ttggttgccggcggtcttatcgaaaactgtgaggctcaactcaatgaacag 2040  
QY 1862 -----aggacctcgggtcgtacgaacaacaagaacatgagaagacc 1902  
Db 2041 aaggaagctgacgcttggaggcgcgcgtcaggtatcagaaatcagatattgccc 2100  
QY 1903 ctgcagaaaaagcagcagagaagtctgtgtgaagacgcgcacttgagcgaactcac 1962  
Db 2101 aaaaacaaactcgtcgaagcgtatccgcatctcagcaggaacacacacacacattgac 2160  
QY 1963 accactgcacatctcaagctggagacctggagaagc-----ggaagggtgcgggtcgg 2016  
Db 2161 cgcacggggcaccacaagctctgttccagcgtattcttctaaagaccccccgaaaca 2220  
QY 2017 actttgggtact-----cactcctaaatgcactcetaatcatgttctcagcaaca 2067  
Db 2221 actcgaataactctatttgacgcgtcctcaaatgtgtcttctcactcaacccgcctgaa 2280  
QY 2068 aaaaactctactgtgtcttctgttgcacaataacaggttctgtatgtgcgggcaca 2127  
Db 2281 atccgtacatcgtcgtcgttctgttctcaggttagctgtgtgtggcacaacccacg 2340  
QY 2128 ttccaggaacaagccctttgttttcgagattgctcactgtcttctcaggttggtgcac 2187  
Db 2341 caagcagtgctatagctaaagcctcagcacactcgtcccgacgaatcacatacagc 2400  
QY 2188 gggcacaaggaacggcgagactttatctgtgttcttctgtgtgtgtgtgtgtgtgtgt 2247  
Db 2401 aggttccagatgacgggatttctgttcattgtgttctgtgtgtgtgtgtgtgtgtgt 2460  
QY 2248 ctatgaggtattcagcaggt 2307  
Db 2461 ttatcagcagcaacaacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2520  
QY 2308 cgtatcaagcgcgaatgttccaaagcagtaactgataagaacatcagcactcgtcagacac 2367  
Db 2521 cgcgcgagaagcactgcttctcggcagcagcactcgcgcgaagcattgttcttggagaag 2580  
QY 2368 cgggagcaaatctcgtgtctcagcgcgaactcgtcagctctaccacacagcagttgcaa 2427  
Db 2581 gaaggaatagcacgcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2640  
QY 2428 gagtgtatcagcaaaatttctcat---ttatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2485  
Db 2641 ggtgttagcgtgtgactcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2700  
QY 2486 agtgcctcactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2545  
Db 2701 atcattattgcctcggcagatgggtggaaatggccttagttgtatcgtgtgtgtgtgtgt 2760  
QY 2546 ccacccctcgttctgt 2605  
Db 2761 gttctcgtggcagcgggttcttaccagattctatagctag---cccagtttcaatcacgc 2817  
QY 2606 aactcggcgaacttgcagaagtgctgggtcgaagcgaagcagttaccgcgagcagc 2665  
Db 2818 tccagctcgtctaggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2877  
QY 2666 accgtctcacttgcactcgaagccatgttctcacaagtaactcagcagcagcagcagcagcagc 2725  
Db 2878 acagttgcgtcattacccgggaagggatgctcgtggagatttaccatgccagcct---tg 2935  
QY 2726 aaggttctgtcagatcattccaaagcctttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2785  
Db 2936 acgcacaagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2995











Db 4618 tatgacaagctgagatcggaatgcgaactctctggtctactgagttctcttatcgctggatt 4677  
Qy 2242 ggaatactagtaggatatattcaagattgctggagcatgcaacgcttatttcaagaattgct 2301  
Db 4678 gtccagatcatcaactctcagtcacaggtatcgctctcggttcagctctgagagattg 4737  
Qy 2302 accatcgtatcaagccgcgaatgttccaaagagtagctggatcaagaacatcgaaactctc 2361  
Db 4738 attcgcaaggtctcgccgaatgccttcctgctgcatgctccgacagagatatcaactcttc 4797  
Qy 2362 gacatcccgagcaaatctctggtctcactcgaactgcaactgctcctccacgcag 2421  
Db 4798 gatcgtgaggagaatagcaactgaggctttgactctcttcttcaactgaacccaagcac 4857  
Qy 2422 ttg--caagagttgatatcagcaaatcttctcaatttatctggttctggtcaacatgctc 2479  
Db 4858 ctggcgggtatcagtggtcaactcttggaaactctctcaactgacttcaacacactcatt 4917  
Qy 2480 tcgagcagtgctctaccactagctatgagaaactggaactggcctgggtgtgtgttgg 2539  
Db 4918 gcttcaattgtagcgcgtgtgttctcgga tggaaagctgtctcgttctgtgatgt--cc 4974  
Qy 2540 goacttcacccctgcttttggctggtcactcagaattctctagagacgaagctagaa 2599  
Db 4975 gtcatcccgattctctcggtcggtctctaccgtttctacatgcttgcgcgttccag 5034  
Qy 2600 gccggaactcggcaaatcttgcaaaagtgtgggttgcaagcgaagagttaccgcg 2659  
Db 5035 gcaagtccaagggcgtctacagaggtacagcaagctacgctcgaggtacatctgccc 5094  
Qy 2660 atccggacgtctcatctttgactctcgagggcctggtctcccaacagtagctcgacatg 2719  
Db 5095 attcgactgtggcttgcttactcgtgagacagctttgtgttgcattctaccatg--gtc 5152  
Qy 2720 ttgagcaaggtcttgtaagatcaatccaaagcttttgggttggagcagtggtttgttcc 2779  
Db 5153 aactcgagctcagggcgcgaagcctgatctcggtattcaaatcatcgtctctgtatg 5212  
Qy 2780 actgtcacagctgagtgagtttggctattgcccctgggatttt--gtattgcagctgat 2838  
Db 5213 ctgcctccagcgttggcttcttcttgggtgcgccttggcttggatgtggaaaccc 5272  
Qy 2839 aattggtctcaggtgagtagacacaactcaatttatatcatcttctgggctgtttgt 2898  
Db 5273 tgcctggccaacatgaatacagacgttctcgtctctcgtgtctcagtgaaatcttgt 5332  
Qy 2899 ttgcgggtccaagcagcagccagttattggtcttactccacagagttttaccaaagctcgg 2958  
Db 5333 tgggtcgtcagctcgtggcagagttctctctt--ctcccgcgacatgggcaagccaag 5391  
Qy 2959 tcggctcggaactatatactctgctcggaacattgaaagccgacacatccggaaacggag 3018  
Db 5392 aacgcgcgcgcagattctctagttgttgagcgtcgccgctacaaattgacacatggctcc 5451  
Qy 3019 gagacaagaataaaggcccgagtggttgatgcctctgcgacctcgaggacattgaattc 3078  
Db 5452 gaagagggcgaataattggaccactggaagaaacattgagttcaaggagcgtcacctc 5511  
Qy 3079 aggtatcgtcaacgtgattcggctcgtcagttctccgcgggggttccatgacaatcgagcca 3138  
Db 5512 cgttaccacccaccctcccgagacgctgtgtccgcgggttgaaacctcactgtcaagcct 5571  
Qy 3139 ggaacaattgtagcttatgtgggcgtctctggtgtggcaagtcaacgttgtatcgctttg 3198  
Db 5572 ggcagtagatctgcctggctcgacactagtggatcgcggaagagtagcacattgcccctc 5631  
Qy 3199 tcggaacgattctacaccgcacctcgggcgaatttctcatttgcacacgagaattatgca 3258  
Db 5632 ctcgaaagtttctacagatgctcttctccggaggtgtctacatagatgacaagaatatcgcc 5691  
Qy 3259 gaaatgtcgcgcgctgtaccgcggcattatgtcttctgttccaaaggaacccacactt 3318

Db 5692 gacctcaatgttaattctactacgcagtcacctggtcgctgggtcagccagagcccccctc 5751  
Qy 3319 taccaaagctcgttcgcgagagaatgtgagcttggccctcgaaagccgaattattcagaagag 3378  
Db 5752 taccagggctacatttaaggaagaacatctctgctcgttagccctcaacaaagatccaacccgag 5811  
Qy 3379 ctttgtcaagagac--gcctccgcgaagcccaatgcttttgagattttgtcatctcttacc 3437  
Db 5812 gagagcgttgcagaggtctgcaaggaagcccaacatctacgacttcatctatgctcgtccc 5871  
Qy 3438 gaagcgtttgaaagccttgcggtcctcaagagggatgcagttctccgcggggcaacgaca 3497  
Db 5872 gaagcgttttaacacgcgtcgtcggaagcaag--ggagcgtgttgcggtggccaaagca 5930  
Qy 3498 gcgagatcccatcgcgaagagcatgttgcgaataatccaaagctgttgcctacttgacgaagc 3557  
Db 5931 gcgtgtcgccatcgccgcgcctcctcgcaatctcaaggtcctacttcttgatgaagc 5990  
Qy 3558 gacgtcagcctcgcacgcgaatcggaacgtctggttcaagctccctcgcgtgaggaac 3617  
Db 5991 tacctcggtctcgaactcgaatccgagaggtcgtccaaagctcgtcgtggtcgtcgtgc 6050  
Qy 3618 cagagccgaacgaataagcagtggtgcacgcacttccacttccggaattgttgatg 3677  
Db 6051 tcggtgtcgtactacatcgctgtggtcacgcgtctcagtaaccattcaaaagccgatat 6110  
Qy 3678 tattttgtttgttccaaaggaagaaatcgccgaacgggacactcagcgggaactacaacg 3737  
Db 6111 catctatgctctcgaacgaagattgtcgagagcgttactcactgactgactgactgcttgcg 6170  
Qy 3738 actgagaggaagatatattacgagatg 3762  
Db 6171 caacaaggtcgtacttctcgagctg 6195

RESULT 8  
AAF83400  
ID AAF83400 standard; DNA; 8777 BP.  
XX  
AC AAF83400;  
XX  
DT 09-JUL-2001 (first entry)  
XX  
DE P. chrysogenum ABC transporter aal0 genomic DNA.  
XX  
KW Beta-lactam; micro-organism; ABC transporter; ATP-binding cassette;  
KW adenosine triphosphate-binding cassette; antibiotic; penicillin;  
KW cephalosporin; aal0; ds.  
XX  
OS Penicillium chrysogenum.  
XX  
FH Key Location/Qualifiers  
FT CDS 2227..6228  
FT FT /\*tag= a  
FT FT /product= "aal0"  
XX  
PN WO200132904-A2.  
XX  
PD 10-MAY-2001.  
XX  
PF 03-NOV-2000; 2000WO-EPI1489.  
XX  
PR 03-NOV-1999; 99EP-0203684.  
PR 03-NOV-1999; 99EP-0203685.  
PR 03-NOV-1999; 99EP-0203687.  
PR 03-NOV-1999; 99EP-0203688.  
PR 03-NOV-1999; 99EP-0203689.  
PR 03-NOV-1999; 99EP-0203690.  
PR 03-NOV-1999; 99EP-0203691.  
PR 03-NOV-1999; 99EP-0203692.  
PR 03-NOV-1999; 99EP-0203693.  
PR 03-NOV-1999; 99EP-0203694.  
XX







Db	2247	acagccacaacattgttccattgctcgcggaagattgccgacacagggaaacacatgac	2306
Qy	1816	gaattgatgaattcggctgtctattacgcccgaactggctgctgcacaggaactctggggt	1875
Db	2307	gaattggtcgatcgcaagcgcgtactataaactctggagcgcaacgtatacaacgag	2366
Qy	1876	gacgaacaacagaagaactgagaaagccctgcacgaagaaggcagcacagagaactctgctgt	1935
Db	2367	gagaagaagcagaagctctggaagccgacgcgcagatggacgcgcatgattttggtcaa	2426
Qy	1936	gaagc-----accggcacttgagcgcactcacacacactgccacatctcaa-----	1980
Db	2427	gaaggggttactcgcatcaagactcgattagcagctcgaaattctctcgatgcttagat	2486
Qy	1981	-----gctggagacctggagaagcggaagtgctccggctcggaactttgggctact	2029
Db	2487	gagaagcgcgctctggagatgaagcgaaacccggaacgcaaaaatcggtttcgagtgccgtt	2546
Qy	2030	cgctctctaaaa-----tgcctctaaatcatgttc	2058
Db	2547	ctatctaagaagaagtcgccgagcagttcgagaagattacctctggaccctcgccaagttc	2606
Qy	2059	tacgaacaaaaaattctactggctcttctgtgttgcacaaataaagggttctgatatgc	2118
Db	2607	atcggtgcatataccgcgcctgagctgtgttacatgctcatgtggcttgactttctogttc	2666
Qy	2119	gcggccacatttcaggacaaagccctttgttttg-----agattgctc	2163
Db	2667	cttgcgggtgggtgtcaacctaccagcgtttctctgtacgccaaggccaattagcacgctc	2726
Qy	2164	actgtctctcgagttgagtggtcatcgccgacaggaacggcgacacttttatattctgatg	2223
Db	2727	tcgttacccgaaatcaatgtttcacaaactcggaagtgacgcgaattctctggtccttgatg	2786
Qy	2224	ttctttgctggtgctctaggaattctagtagatatattcaagatttggctggacatgcaac	2283
Db	2787	ttcttctggttggaaattgctcaatttatcagcgtgtctatacaaatggtcagcatttgc	2846
Qy	2284	gttatctcaaaattgtctaccatcgctatcaagcgcgaatgtttccaacgagttactggtat	2343
Db	2847	attgttcggagagactcaattccgcagctagaagtcgaacgtttagatcgattcttcgt	2906
Qy	2344	caagacatcgaaactctcgaaatcccgagacaaattcttggtgctctcaatcgcaactg	2403
Db	2907	caggacatctatttttggacagggaagagaaacgacgacgggtgctctgacgtccttcta	2966
Qy	2404	tcagctctaccacgcagttgcgaaggttgatacagcaaatcttctcatattatctgtt	2463
Db	2967	tcaacagagacgaagaatctgctcgggcgttagcgagtgactctggcacaatcatcatg	3026
Qy	2464	gtcgg--tcaacatcgctctcgagaagctctataccactagcctatgatggaaactggg	2521
Db	3027	actagcaactacgcttgggcgcgcatgatcatgttcattggtcgatcggaactggagctg	3086
Qy	2522	ctgggtgtgtgtttgtgcacttccacccctgttttggctgctaccctcagaattcgt	2581
Db	3087	ctgggttgcatttctgtcg---ttccaatctctctgcatcgcgctctctcagattctac	3143
Qy	2582	ctagagacgaagctagaagccggaacctcggcaactttcgaaaaagtgtcggttgca	2641
Db	3144	atgcttgctcaattccagcaacgattcgaaagtctgcctacgaaggggtctcgagactatg	3203
Qy	2642	agcgaagcagttaccgcgcatccgacgctctcatcttttgactctcgaagccaatgtctc	2701
Db	3204	tgcgaagccactcagcatccgcactgtagcatcaactcactcgcgacaagaatgtctcg	3263
Qy	2702	caacagtactcggacatttgagaaggtcttctaagatcatccaaagcttttggtttg	2761
Db	3264	ggcgtttcacgcgacacgt--acaaaaacagggaagcgaagtttgatctcagttctga	3321
Qy	2762	gacgatgtttgttttcaactgttcacagtcgattggagtttttggctatgtccctgggat	2821

Db	3322	gatctccctgctgatatgctcctgctgcagggcattggtgttttctctgcgcgtctttgggct	3381
Qy	2822	tttgtattgcagtcgataaattggcttcaggtgagtcagacacaactcaatttatcat	2881
Db	3382	tctggtatggtgtacactctaggccatcaatgagtacagcatctccgctctctcgtct	3441
Qy	2882	cttcgtggcggttttgtttgctcggtctcaagcagacgcccaagtatttggtcttaactcacga	2941
Db	3442	gctttctcgagatctttttgtgtgcacatcagccggaactgtctctctcttttgcgccgg	3501
Qy	2942	gtttaccaaagctcgtcggtcggaactatactcttgctggcggaacttggaagccga	3001
Db	3502	acatgggtaaggcaagaatcgcgctgctcaattcaagaaactcttcgacagcagaccaa	3561
Qy	3002	ccatccgcgaacgcgagggagacagaaaaagggccagtggtggtgccctgctgcgacc	3061
Db	3562	ccattgacatctggtcggaatgagggcgagaagttgagttctatggaaaggcgaatcgat	3621
Qy	3062	tcgaggacattgaattcacagttatgctcaacgtgattcgcgtcgagttctccgcgggttt	3121
Db	3622	tcgggacgtccactttagttacccaacggcgcgagcagcctgtttctcgaggactga	3681
Qy	3122	ccatgacaaatcgagcagacgaataattgtagcttatgtgggcgtctctgctgtgcgaagt	3181
Db	3682	atttgagcgtgaagcctggacatacatctgccttgttggaocccagttgagtcgtaaga	3741
Qy	3182	caacgttgatcgctttgtcggaacgattctagaccgcgacctcggcgcaaatcttatgtt	3241
Db	3742	gcactacgattgctcgtcttgagcattttatgacgcacttgctggaggggtcttcgttg	3801
Qy	3242	caacgagaatatcgagaaaatgtcgcgcgcgtgttgaccgcggccatatgtcttttgttcc	3301
Db	3802	acggaaaagacattaccaaaccaatgtcaactcatcacgcgagttctctcctctgtca	3861
Qy	3302	aacaggaaaccacactttaccaggctcgtgtcgcgagaaatgtgacgtttggccctcgag	3361
Db	3862	gccagaacctactctgtatcagggtaccatcaaggaaaaatatcctgctggagtcgata	3921
Qy	3362	ccgaattatcagaagagctttgtc-aaggacgctctccgcgaaggccaatgctttggatt	3420
Db	3922	aggatgacgttttcgaggagacatttgattaaaggctcgcgaagatgccacatctatgatt	3981
Qy	3421	tgtcatctctttaccagaag-----	3441
Db	3982	tcgttatgtcactccgtgaagttaactgctcgtgactcctcttcatgctgcggcggtct	4041
Qy	3442	-----gctttgaaacgcttgcggtccacagaggatgca-gtt	3479
Db	4042	gctaacatgctcacacagtgaaggatttgacacgcgtgtgacagcaaggaggcattgtt	4101
Qy	3480	ctccgcggggcaacgacagcggatcgccatctgcaagagcatctgattcgaaaatcccaagct	3539
Db	4102	gtctgttgacaaaaacagcgtgtgcgcattgctcgtgcctcctcgtcgtgaccccaaggt	4161
Qy	3540	gttgtactttgacgaagcagactagcctcgcacgcgaatcgggaacgctctggttcaagc	3599
Db	4162	ccctctcttgtagtaggcacatctgctctgactccgaattctgagaagtcgtacaaagc	4221
Qy	3600	tgcctctgatgaggtatccacgagcgcgaacgacatagcagtggcgcacacgactttccac	3659
Db	4222	tgcgtggaatgctgtgcgcgcggcggaacacgattgctgttgccacacgcggctgagcac	4281
Qy	3660	tattcggaaatgttgatgtatttttgttttcgaacggagaaatcgccgaacagggcac	3719
Db	4282	ctttcaaacgctgataaattatctgctctcgacaaaggcagagctgctcggaatgggaac	4341
Qy	3720	tcacgcgaactcacacgactgagagggaatattacgagatg	3762
Db	4342	gcaccacaaatgtattctaaacaaaggccggtattcacagctg	4384

RESULT 10  
AAT35230









Db 1831 gaaaaatttgactgtggaagaggcttgtgtogattgcacaaagccataattttatcatg 1890  
QY 1519 gagcttccgaacgggttatgagactgaagttggcgagcgagcgagccttgagtgaggt 1578  
Db 1891 aaacttccgaaggttatgataccatggttggtagcgcggtatgctgttatctggtggt 1950  
QY 1579 caacgacacgaattgcaatcgcgaagaatgatatcatcgtgatcccaagatcctgttactc 1638  
Db 1951 cagaagcagcgagtgctgattgctgtctgtatcgtttccgacctaggatttggttg 2010  
QY 1639 gatgaagctaccagcgcccttgaccgaagggcggaagaagtgggtccagagggccttgaac 1698  
Db 2011 gacgaagctacactcgtcttgacactcagagcggaaggtattgtgcaaatgctttggac 2070  
QY 1699 cgagtgctcaaaagaccgcactacttggctcattggcccacaaactagccactgtcaaaagt 1758  
Db 2071 aaggctcgaggcaggaactactatcaccaattgctataggttatctaccactcgagat 2130  
QY 1759 gctggcaacatcgagctatcttcccaaggggaaaaatcgtcgacaaagggcacacaccacgaa 1818  
Db 2131 gccgacagaatttaagtgtggtggtggtggcggaagtactggagcaggggttctcacaacgat 2190  
QY 1819 ttgactgaattcg---gctgtcattacgcgcactggtgctgtgcacaggaacctcgggct 1875  
Db 2191 cctctgcgaacgaaacggctctctacgcacaactgtcaacaactgtcaacaacttgcacaa 2250  
QY 1876 gacgaacaacaagaacatgagaaagacctgcacgaaaaggcgacgacgagagctgctggt 1935  
Db 2251 gaagctgctgtgaagccctcaggtggatgacgatatcgaggaacctgacgatgcggt 2310  
QY 1936 gaacgacggcacttgagcgactcacacactgcccacatcgcacatctcaagctggagacctggag 1995  
Db 2311 ttcatggcggttcttcgctatgcgaagaaaagagacacacagcttcatcgcgagtgact 2370  
QY 1996 aagcggaaggtcggtcggtggttggctactcgctcctctaaatgcactcctaactatcag 2055  
Db 2371 ggaaggtctcttgccagtatcgccatggacgacatccagccgaagggcgagggaagtc 2430  
QY 2056 ttctcgaacaaaataatctctactggtgctctctgttctgttgtaacaataacgggtctcgata 2115  
Db 2431 gcaggtgaggacaagataaccttcaagcttggctgtgatgccaggtgttgagatgaat 2490  
QY 2116 tgccgggccacatt----- 2129  
Db 2491 tcgcgcgataaattatatacatatagcgttttatcgccgcacatctgcgagcgatggtt 2550  
QY 2130 ---tcgagacaagcccttgttttcgagattgctcactgtcttcgagttgagtgatcat 2187  
Db 2551 tatccttcttggccatcctcttggcaagccctctccgacttggagattcaagatcct 2610  
QY 2188 ggggcacaggaacgggcagacttttatatt---ctgattcttcttgcgtggctctagga 2244  
Db 2611 gccgaactcgacatgctctcagtcgcagtgccctgtggtactttatcactgctcttg 2670  
QY 2245 aatcagtaggatatattcaagattggctggaatgcacaagttattttcaaaagtgtcacc 2304  
Db 2671 gctgattgtcatctattcttccaaagtgcggctctctctcgcgcgggttgggatttgaac 2730  
QY 2305 catcgtatcaagcgcaattgttccaaagagtactggatcaagacatgaactcctccgac 2364  
Db 2731 ggtgtcttgagaaaaaagtgttctactgctacttggacatgacatcgaatgggttcgac 2790  
QY 2365 atcccgagacaaattctggtgctctcacatcgcaactgcactgcactaccacgcagttg 2424  
Db 2791 gaggaaacaaactctacttgggtccgttcacatccaaaccttgcgatcaacctcagaaggtg 2850  
QY 2425 caagagttgatacagaanaatttc---tcatttatatggttgcgtgcgaactatgctctcg 2482  
Db 2851 caaggtcgttggcctacttcttgggttcagctgcgtccagctcgtgcgcaactctcatcgt 2910  
QY 2483 agcagtgctctaccactagcctatgtaggaactgggacctggtgtgtgtgtgtgtgca 2542  
Db 2911 ggttgattatcggtctgtgtaaggtctctctctctgctctgctggtatcggtcttgt--- 2967

QY 2543 ctccacccctgcttggctggctacccagaaattcgtctagagcagacgaagctagaagcc 2602  
Db 2568 atcccactcctgtctctggaggttatattcgttcaagggtcgttactcaaggaacaa 3027  
QY 2603 ggaacctcgcaaacctttgcagaagtgctggcttgcgaagcgaagcagttaccgcgactc 2662  
Db 3028 cgtatgaagaagctccatgcgcgtagtgacatcttgcagtgaaagctgctgagccgct 3087  
QY 2663 cggacctctcatcttggactctcgaagggccatgttctccaaagctactcgaacatgttg 2722  
Db 3088 aadacagtagccagctcttaccagggagaaggtatgaaggagaatctatctgaaagctctt 3147  
QY 2723 agcaagctcttgcataagatcatccaaagcttgggttggcagatgttgggttttctact 2782  
Db 3148 a--aagctctatgaataataattccgtactcttatataatccacagtgctctcttggcg 3205  
QY 2783 tgcacagtcgatggagtttttggctattgcccgggattttgttatt--gcagtcgataat 2841  
Db 3206 cgagtcaggggttgacttctgtatcattgcttgcgttctctacattggcgattatgga 3265  
QY 2842 tggcttcaggtgagtagcacacaactcaatttatatatctctctgtggcgcttggttg 2901  
Db 3266 tcactgatgcataactcactgctggttctctactgcttctgaaactccatctgctcttg 3325  
QY 2902 ccggttccaaagcagcagcccgatattggcttacttccacagagtttaccaaagctcggctg 2961  
Db 3326 -cgtctattcaagctggaatgtctctcaccttgcgcggtatgggtcccaagccaatctcc 3384  
QY 2962 gctgcgaactatatacctctggctgcgacatctgaagccgacacatccgcga-----aacg 3015  
Db 3385 tcggcgccctccacttctccgaagcatcgacaatgaacggctatcaatgctgaaactaat 3444  
QY 3016 gaggaaacaagaanaaaagccagtggtgagtcctcgcagctcgcagcgaagacattgaa 3075  
Db 3445 gaaggggaagtgcttgatccaaagcagctgctggtcatgtggaattggaagggctccac 3504  
QY 3076 ttcaggttctgctcaacgtgattcggctcaggttctccgcgggtttccatgacaaacagag 3135  
Db 3505 ttcaagttactaccagcgcggtgagtagtactgaggaatttgaactatgatgttct 3564  
QY 3136 ccaggacaaattttagcttatgtggcgcttctgctgtggaagtcgaagtcgaactgactc 3195  
Db 3565 gccggcaactatgtggcactcgttggctcctcggttgggaagttaccactactcaa 3624  
QY 3196 ttctcgaaacgatctacgaccgacactcgggcgaattctcatitgcacagagaaatatt 3255  
Db 3625 atcgtggagaggttctacgaccttggctggcggtgacttggatggatggatgatatt 3684  
QY 3256 gcagaaatgtcgcgcgcttgtaccgcggccatattgtcttgggtccaaacaggaacccaca 3315  
Db 3685 aaggagctcaacttggcgagttaccgtagccagatctccctcgtttctcaagaaacctacc 3744  
QY 3316 ctttaccaaaggtccttgcgcgagaatgtgacgttggccctcgcgaagcgaattatcagaa 3375  
Db 3745 cttttacggggttcaaattagtttaattcttcttggcgcaacaaagccctcagagagag 3804  
QY 3376 gagcttctgaaggacgctctccgcg-----aaggccaatgctttgagatttctgctac 3428  
Db 3805 gtactcaggatgagattgatgtcgtgcgaaggtgccaaatattacgaacttatctg 3864  
QY 3429 tcttaccagaagcctttgaaacgcttgcggctcaacgagggatgcagttctcccggtg 3488  
Db 3865 tcttgcagatggttttgacac--cgaaigtgggtgggaaggtctcagctctctcgttctggtg 3923  
QY 3489 gaaacgacagcgatcgccatcgcaagagcatgtatctcgaaatccaaagctgtgtactc 3548  
Db 3924 tcaaacgacgtatcgctatcgccgtgtctctcatcgaaacccctaaagtctgtctct 3983  
QY 3549 tgacgaagcgacgtcagccctcgacacgcaatcggaacgtctgtgttcaagctgcctcga 3608  
Db 3984 tgacgaagctacgtccgcactcgacatcagtcgaagaaaggtctgttccagggagcacttga 4043







```

RESULT 14
AAF83381
ID AAF83381 standard; DNA; 7744 BP.
XX
XX
AC AAF83381;
XX
DT 09-JUL-2001 (first entry)
XX
DE P. chrysogenum ABC transporter aal genomic DNA.
XX
XX
KW Beta-lactam; micro-organism; ABC transporter; ATP-binding cassette;
KW adenosine triphosphate-binding cassette; antibiotic; penicillin;
KW cephalosporin; aal; ds.
XX
OS Penicillium chrysogenum.
XX
FH Key Location/Qualifiers
CDS 1383..5674
FT /*tag= a
FT /product= "aal"
FT /note= "contains introns"
FT 1383..1696
FT /*tag= b
FT /number= "1"
FT 1697..1749
FT /*tag= c
FT /number= "1"
FT 1750..2406
FT /*tag= d
FT /number= "2"
FT 2407..2461
FT /*tag= e
FT /number= "2"
FT 2462..3639
FT /*tag= f
FT /number= "3"
FT 3640..3695
FT /*tag= g
FT /number= "3"
FT 3696..3871
FT /*tag= h
FT /number= "4"
FT 3872..3924
FT /*tag= i
FT /number= "4"
FT 3925..4028
FT /*tag= j
FT /number= "5"
FT 4029..4093
FT /*tag= k
FT /number= "5"
FT 4094..4454
FT /*tag= l
FT /number= "6"
FT 4455..4539
FT /*tag= m
FT /number= "6"
FT 4540..4690
FT /*tag= n
FT /number= "7"
FT 4691..4769
FT /*tag= o
FT /number= "7"
FT 4770..4901
FT /*tag= p
FT /number= "8"
FT 4902..4999
FT /*tag= q
FT /number= "8"
FT 5000..5671
FT /*tag= r

```

```

FT /number= "9"
XX WO200132904-A2.
XX PD 10-MAY-2001.
XX PF 03-NOV-2000; 2000WO-EP11489.
XX PR 03-NOV-1999; 99EP-0203684.
XX PR 03-NOV-1999; 99EP-0203685.
XX PR 03-NOV-1999; 99EP-0203687.
XX PR 03-NOV-1999; 99EP-0203688.
XX PR 03-NOV-1999; 99EP-0203689.
XX PR 03-NOV-1999; 99EP-0203690.
XX PR 03-NOV-1999; 99EP-0203691.
XX PR 03-NOV-1999; 99EP-0203692.
XX PR 03-NOV-1999; 99EP-0203693.
XX PR 03-NOV-1999; 99EP-0203694.
XX (STAM ) DSM NV.
XX Van Den Berg MA, Bovenberg RAL, Driessen AJM, Konings WN;
PI Schuurs TA, Nieboer M, Westerlaken I;
XX P-PSDB; AAB62488.
XX WPI; 2001-291055/30.
XX P-PSDB; AAB62488.
XX Enhancing secretion of beta-lactam compounds from a micro-organism by
XX enhancing adenosine triphosphate-binding cassette transporter activity,
XX useful for producing e.g. penicillin and cephalosporins -
XX Claim 9; Page 44-56; 116pp; English.
XX The invention relates to a method for enhancing the secretion of beta-
XX lactam compounds from a micro-organism that comprises enhancing ABC
XX (adenosine triphosphate (ATP)-binding cassette) transporter activity of
XX the micro-organism. The method is used for enhancing the production and
XX secretions of beta-lactam antibiotics like penicillin and cephalosporins.
XX Manipulation of the ABC-transport protein system provides a means for
XX enhancing beta-lactam secretion. The present sequence represents the
XX genomic DNA sequence of the P. chrysogenum ABC transporter aal.
XX Sequence 7744 BP; 1960 A; 1847 C; 1934 G; 2003 T; 0 other;
SQ
Query Match 8.48; Score 319.8; DB 22; Length 7744;
Best Local Similarity 51.2%; Pred. No. 3.8e-89;
Matches 843; Conservative 2; Mismatches 744; Indels 58; Gaps 2;
QY 295 ctgtgggtcgtctacctctttattgggaagtgttgcctcacgtacgtttccacggctgcc 354
Db || || || || || || || || || || || || || || || || || || || || ||
Db 1751 ctctactcttctgtcattggcctttgcgattcttttaactttatggatagacacactg 1810
QY 355 attaccatttcagctatcacgaaccacgtcgaactcttcgacgagtgcttcctgaatgcacc 414
Db || || || || || || || || || || || || || || || || || || || || ||
Db 1811 tttaacaattgcgcgtaccggtacgcgaaataattgcgccctgtatctgaaagcaggt 1870
QY 415 ttgcggcaagaggtctgtggaatttcgacacagcagcaatggagcaatcgccactcargtc 474
Db || || || || || || || || || || || || || || || || || || || || ||
Db 1871 ctgaccacaggaaattgcctctcttgcgtgtgctggtggtggtggtggtggtggtggtg 1930
QY 475 actaccaatggcaaccgtatatacaaacagggtatttcgcgagaaattggtttaccgtgcag 534
Db || || || || || || || || || || || || || || || || || || || || ||
Db 1931 acctcgaacggacgattgattccagagtggaataatcggaagagctggcctgtgtccag 1990
QY 535 gcaatttcaatgtcttcttctgtcatttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 594
Db || || || || || || || || || || || || || || || || || || || || ||
Db 1991 ggctgtctactttatataccgcctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2050
QY 595 ttaatcaccatgtccgtcatcctcgtcatttctgttcaccgcatcgtcatgataat 654
Db || || || || || || || || || || || || || || || || || || || || ||
Db 2051 ctoatttgcgtgtattgcccacagcaattctctcgcctctaggaatcagctcgtgcatt 2110

```



Db 348 gatlgacttggtgactgcttgtagagcggaataatcatatcagagctgaogctgatttc 407  
Qy 261 tgaagcattcaagccgagatggatcatttcgctcgtggttgcgtacacgtctttattgg 320  
Db 408 taacttgcttgagcaaaagtgcgaggttctccctgcaaaacccacacatcgacacat 467  
Qy 321 gaagtgtgctcaactgactgttccacaggtgccaattaccatttcaagctatcagaaccac 380  
Db 468 tatgctggtgttcctatctctcgatcactgcttcaactacggtcccaactccagat 527  
Qy 381 tcgaactcttcgacagagtggttcttgatgaacaccttgccgcaagaggttggaatttcga 440  
Db 528 cctcacatccgctcaaaagttcttctgcctccattctcgaacgacatgaagtgggtacga 587  
Qy 441 caagcagagcaatgagcaatcgcacactcargtcactaccaatggcaacgtatacaaac 500  
Db 588 ctccaaccagagtggcgaagtagccagccgaatgaatgaagatctccaaaatggagga 647  
Qy 501 aggtattgcgagaaatggtcttcttacggtgcagagcaatttcaatggtttttctgcat 560  
Db 648 tggcttgccgaaaaagtggtcattgctccatttacccttgctccttcgtgggatcttt 707  
Qy 561 tgtgctgcttggtcgtcactgagtggaagctagcttttaaccacatgctcgtcatccctgc 620  
Db 708 ggtgctagcctttgcaagggtgcgagctatccctggtgctgtgctgaccagtttgccact 767  
Qy 621 cattttcctggtccaccgactctcctatagcaaatgtagccgctcagagagccagagatacc 680  
Db 768 aacattcatcgccatgggctggtggcgtggccacgtccgcctgctcgaaaaaggaggt 827  
Qy 681 caggatctactcacgcgcgctgctcctcgagaaagagctcttataccatccggagacgt 740  
Db 828 gaccatgatgctggtcgtccgctgggtgagtgagcgctctcctccggaattcgaaacgt 887  
Qy 741 ccatgctttctacgcccagaanaaaatggtcgaaaaatatgctctttttgagcaagc 800  
Db 888 aaagcttttgaggcgagcgaaagggtggctgcgtacaaagagcggtggtgtgctgc 947  
Qy 801 acaccaagaaggaagaagaaatcgccaanaataatggsgcttcttcaactgagactt 860  
Db 948 caagatcttgaaatacaagaaataatgtctccggaatcggaattcggtctgcttggct 1007  
Qy 861 ttgcatcttaoigtctatcgcaactggccttttgaaaggttttgcgcatgatacagaatgg 920  
Db 1008 ctctcatcagcactcctatgctctggccttctggtatggagtggtggtggaatcaagg 1067  
Qy 921 -----cgaggttgccgacgttggcaagctcttactgctgctcttc 962  
Db 1068 ataccagaaaccggtatgagaactacgtagctggcaccatgataccgctattcttcttc 1127  
Qy 963 cgtcacctttagcagccgctcatcgaatgcttgcgcttcaaggttcaggttcgatttacc 1022  
Db 1128 cgtgatgatgggttccatgaataatggaatggcgctccctatatacgaggcatttggcat 1187  
Qy 1023 acgcgcgcatcttgctccgaattattcagtatcatattgacaaaacccagcagctgcacc 1082  
Db 1188 cgcgaagggcgccctgcgccaaggttttccacattatgagcagattccggaaattatcc 1247  
Qy 1083 tctcgaccccttctggaagcagcagagggctgctgactagggtcaaatggatgccaaaacct 1142  
Db 1248 cattgatggggagggtaaaaagttaaatgagcccttaccacattgagtttaagggaagt 1307  
Qy 1143 ggaatttgcctaccctccgacacatctgcgaagtaacttcgagatttcaacttgacaat 1202  
Db 1308 ggaatttccagtatcccgacgcgaggggtttctctgaacaaaactgaaacctgaaat 1367  
Qy 1203 tccagctggcgaagcagcagccctcgtcgtggtcactgggtgagcggaagcaaatgg 1262  
Db 1368 ccatcgctggccagcagcggtggctctgtgtggctccatcgcgctgtggcaaatcaacctgcat 1427  
Qy 1263 cggcttacttgaaacgggtgggtatctgcccagttcgaggaggtatatttaattgaggttgcga 1322

Db 1428 ccaactggttcagcgattcttatgacccccaggctggttaattcttttgcataatggcacaa 1487  
Qy 1323 actgggacatacaatgtgaaatgggtgagaaagccgctacgctcgtcttcaacagggaacc 1382  
Db 1488 tcttaaggacctcgacatacaactggtctacgctctagatggagtggtggttcagggaacc 1547  
Qy 1383 tgtgtgtgttcgtggcaaatcttccagaaacattgccaacgggtttctatgtagagcaaacg 1442  
Db 1548 aatctcttcgcaacttccatatacagagaatactaggtacggttcgagagagcgaactcg 1607  
Qy 1443 agatctcctcgcgaaaaacaaatggagcttgtagaaaaagcttgcaaaagcgaactagg 1502  
Db 1608 cga-----ggagatcgagggcagcagctgcccagc----caatgc 1643  
Qy 1503 cgacgtgttcttaatgagcttccgaacgggtttatgagactgaagtttggcgagcgagcg 1562  
Db 1644 agccattcttataagagctacccaaggatatacttctggtgggagcgagcgagc 1703  
Qy 1563 agccttgagtgagggtcaacgcaacgaattgcaatcgacgaagtgatctatcggtacc 1622  
Db 1704 tcaagtgtccggggggcgaaagcagaggaatagccattgctcgcgcatgtattctgatacc 1763  
Qy 1623 caagatcctgttactcgatgaagctacccgagcgccttgacccgaagcgagaaagtggt 1682  
Db 1764 cgaaatctgtgttgatgaagccacatctgcatttgatataccgccagttaggctaaagt 1823  
Qy 1683 ccaggagccttgaaacgagtgccaagaacccgactacttggctcatttgcacacaaact 1742  
Db 1824 acaagctgcctggagaaagtcagcgctggaaggacacatcatcttggcccaccgttt 1883  
Qy 1743 agccctgttczaaagtgctggcaacatcgagctcatcttccagggggaataatcgctgagca 1802  
Db 1884 gtccacgcttctgctcgggcagacagagattggtggtatcaacaaggttgagggtggagag 1943  
Qy 1803 aggcacacacacgaattgactgcaattcgctgctcatttaccgcsactggt----- 1853  
Db 1944 tgaacgcatcaagagtgtagtgaggttgagggaacacatttcaatttgaacacactca 2003  
Qy 1854 gcgtgcacagacccctcggtgtagcgaacacaaagaacatgaggaagccctgcgcgaaaa 1913  
Db 2004 actgggtgagagtgatggtcctggtcccttagtccacactggtgatatcacaagaacttga 2063  
Qy 1914 ggcagcagagaagtgctggtgaacacgcgcacttgagcgcaactcacacactgcccac 1973  
Db 2064 catcaaggaatgaagatgaggaagagatcaaaagtactcagcgagacgagacgaagatgt 2123  
Qy 1974 atctcaagctggagaccttgagaaagcggaaggtgcccgttgggacttgggctactcgct 2033  
Db 2124 gatggtcactgacgagaaaaataaagaaaaagaaagagaggtcaaggaccccaacga 2183  
Qy 2034 cctaaaaatgcatcctcatgcttcttacgaaacaaaaaatctctactggtgcttctgtt 2093  
Db 2184 ggtgaaccccatgctggaggtgtagaaaaatgaacaaacccggaatgggttcagatcgca 2243  
Qy 2094 gtcaacaaataacgggttctgatatcgcgccacatttccaggaacagaccccttggtttc 2153  
Db 2244 ggggtgcatctcctcgatcatgggtcgccatgccccatttggcgtgctcttgg 2303  
Qy 2154 gagattgctcaactgctctcgaagtggagtgctatcgcgccaggaacggcgagactttta 2213  
Db 2304 tagcatccttcagatttctcgtttaagggaataatgatacaatcgttcgcgaaaaactccaa 2363  
Qy 2214 tattctgatttcttctggtgggtcttaggaanaatctaggtgatttctcacgattggctg 2273  
Db 2364 ccagtaacaggttgtaacttcttgatggcggtatcggttagtggtgggaattgccaggttccata 2423  
Qy 2274 gacatgcaacggttatttcaacagttg-----tcacccatcgctcatcaagccgcaatggt 2327  
Db 2424 gatatacttcttggtagtcggtgagagactgacggaaccccttcgaggaactcatggt 2493  
Qy 2328 ccaacgagtgactggatcaagacatcgaaactcctcgcacatcccgaggaacaaattcttgggtgc 2387  
Db 2484 cgaggccatgctgcgcgagaggtggcctggttcgtagacaagcggaacggttacgggaag 2543





-----

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2002, 23:09:34 ; Search time 7811.87 Seconds  
(without alignments)  
10158.058 Million cell updates/sec

Title: US-09-882-694A-10  
Perfect score: 3792  
Sequence: 1 atggcagatgaatcgagaa.....aatctttagaccagcatga 3792

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_to.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	3791.2	100.0	3792	6	AX076851	Sequence
2	3454.8	91.1	3999	6	AX076850	Sequence
3	543.8	14.3	6120	8	AF071410	Aspergill
4	487	12.8	4002	6	ARI49776	Sequence
5	487	12.8	4002	6	ARI49777	Sequence
6	450.6	11.9	4050	8	SFU62933	Aspergillus
7	424	11.2	4569	8	AF173826	Emericell
8	424	11.2	6133	8	AF071411	Aspergill
9	401	10.6	4647	8	AF375879	Venturia
10	392	10.3	4576	8	AF291822	Trichophy
11	367.4	9.7	5100	8	SFU62934	Aspergillus
12	362	9.5	3924	6	AR013812	Sequence
13	362	9.5	3924	8	AFU62931	Aspergillus
14	351.2	9.3	4227	8	FNU62929	Pilobasidie
15	343	9.0	4872	8	AF082072	Emericell
16	342.4	9.0	35369	8	SPCC663	AL031307 S.pombe c
17	337.6	8.9	4762	8	YSPMD1	DI0695 S.pombe pmd
18	324	8.5	4280	10	CRUPGPII	M60041 C.griseus p
19	306.2	8.1	4175	3	AF003908	AF003908 Haemochu
20	301.8	8.0	3987	10	CRUPGP1185	M59254 Chinese ham
21	301.8	8.0	4296	10	CRUPGPI	M60040 C.griseus p
22	300.2	8.0	4304	10	CRUPGP1165	M59253 Chinese ham
23	300.2	7.9	4788	6	AX322793	Sequence
24	300.2	7.9	4924	10	MUSMDRIA	M33581 Mouse p-gly
25	298.6	7.9	4356	10	MUSMDRIA	M30597 Mouse multil
26	295.4	7.8	4927	10	AF257746	AF257746 Rattus no
27	291.6	7.7	3858	4	OA078609	U78609 Ovis aries
28	291.2	7.7	4323	10	AF286167	AF286167 Rattus no
29	291	7.7	3924	6	AX024455	AX024455 Sequence
30	291	7.7	3924	9	HUMMDR3	M23234 Human membr
31	290.4	7.7	4189	6	AX322791	AX322791 Sequence
32	290.4	7.7	4298	10	MUSMDR	M14757 Mouse multil
33	286.4	7.6	3912	10	RATPGLYCO	L15079 Rat p-glyco
34	286	7.5	3682	10	MUSPGLY1A	M24417 Mouse phosph
35	286	7.5	4233	6	ARI23273	Sequence
36	285.8	7.5	4900	8	AFU62932	U62932 Aspergillus
37	283	7.5	4045	12	AF269224	AF269224 Synthetic
38	283	7.5	4279	6	AX105057	AX105057 Sequence
39	283	7.5	4279	6	AX105078	AX105078 Sequence
40	283	7.5	4279	6	AX105080	AX105080 Sequence
41	282.6	7.5	4317	4	AF045016	AF045016 Canis fam
42	282.6	7.5	4317	6	AX105059	AX105059 Sequence
43	281.4	7.4	4279	6	AX105082	AX105082 Sequence
44	278.2	7.3	3934	4	CFA419568	AJ419568 Canis fam
45	277.4	7.3	4186	6	AX108654	AX108654 Sequence

ALIGNMENTS

RESULT	1	AX076851	Sequence 10 from Patent WO0105980.	3792 bp	DNA	linear	PAT 22-FEB-2001
AX076851	LOCUS	AX076851	Sequence 10 from Patent WO0105980.				
	DEFINITION	AX076851					
	ACCESSION	AX076851					
	VERSION	AX076851.1	GI:13121523				
	KEYWORDS						
	SOURCE	Exophiala spinifera.					
	ORGANISM	Exophiala spinifera					
		Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothyrionmycetes; Chaetothyriales; Herpotrichiellaceae; Exophiala.					
	REFERENCE	1 (bases 1 to 3792)					
	AUTHORS	Duvick J.P., Maddox J., Gilliam J., Folkerts O. and Crasta O.R.					
	TITLE	Compositions and methods for fumonisin detoxification					
	JOURNAL	Patent: WO 0105980-A 10 25-JAN-2001.					
		Pioneer Hi-Bred International Inc. (US) ; Curagen Corporation (US)					
	FEATURES	Location/Qualifiers					
	source	1..3792					
		/organism="Exophiala spinifera"					
		/db_xref="taxon:91928"					
		/note="p-glycoprotein, coding sequence"					
	BASE COUNT	939 a	928 g	940 t			2 others

## ORIGIN

Query Match 100.0%; Score 3791.2; DB 6; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcagatgaatcggagaaaccccgacaaacagatggcagtgagtcgtctctcacac 60  
Db 1 ATGGCAGATGAATCGGAGAAACCTCGACCAACCAAGATGGCAGTGAAGTGGTCTCACAC 60

QY 61 cctccccgaaagaaacccgagagcagatttccagactatctacgaatat 120  
Db 61 CCTCCCCGAAAGAAACCGAAGACAGTATTTTCAGACTATCTACGAATCTTCAGATAT 120

QY 121 ggcgacaaatagactggactctcaatgtcatcgctcatctgcccactcggatccggg 180  
Db 121 GCGGACAAATACACTGGACTCTCAATGTCAATCGCGCTCATCTGCGGCATCGGATCGGG 180

QY 181 gctccctctctctgatgtcgatcatctctcgttagcttccaccacaagttccaacaattac 240  
Db 181 GCTTCCCTTCTCTGATGTGATCATCTTCGGTAGCTTCACCAACAAGTTCAACAATTAC 240

QY 241 aattcgggagcggagtcctgaagcgttcaaggccgagatgtggaatcttctcgttgg 300  
Db 241 AATTGGGGGACGGAGTCTTGAAGCGTTCAAGCCGATGTGGATCATTTTCGTCCTGTGG 300

QY 301 ttctctactctcttatgggaagtttgcctcacgtacgtttccacggtgccattacc 360  
Db 301 TTCTGCTACTCTTATATGGGAAGTTTGCTCACGTACGTTTCCACGGCTGCCATTACC 360

QY 361 attcagctatcgaaccactcgaactcttcgacgagtttcccttgaatgcaccttgcgg 420  
Db 361 ATTTGAGCTATACGAACCACTCGAATCTTCGAGGAGTGTCTTGAATGCACCTTGGCG 420

QY 421 caagaggtcgcatttcgaagcagagcaatggagcaatcgccactcargtcaatacc 480  
Db 421 CAAGAGGCTGGCAATTCACAAAGCAGAGCAATGGAGCAATCGCCACTCARGTCACTACC 480

QY 481 aatggcaaccgtatatacaaacggtattgccgagaaattggctttaccgtcaggcacctt 540  
Db 481 AATGGCAACCGTATACAAACAGGTATTGCCGAGAAATTTGGTCTTACCGTGCAGGCACCT 540

QY 541 tcaatgtttctctgcatttggctgtgttggcttcagtggaagtagctttaate 600  
Db 541 TCAATGTCTTTCTCTGCAATTTGGTGTGCTTTGGGCTCTCAGTGGAAAGTAGCTTTAATC 600

QY 601 accatgtcgtcatccctgcctatttctcgttgcacggtcatctgcatagcaattgatgcc 660  
Db 601 ACCATGTCTGTATCCCTGCTCATTTTCTGCTACCGGCATCTGCATAGCAATTTGATGCC 660

QY 661 gctcaggaggccagatcaccaaggatctactcacgcgcgctgtcctcgcagaagaagtc 720  
Db 661 GCTCAGGAGGCCAGGATCACCAAGGATCTACTCACGCGCGCTCTCTCCGACAGAAAGTC 720

QY 721 ttaatacatcccgagagtcactgttctctacgcccagaagaataatgtcgaataat 780  
Db 721 TTATCATCATCCGGACAGTGCATCTTTCTACGCCCAGAGAAATATGTCGAAAAATAT 780

QY 781 gatgtcttttgcagcaagcaccaagaaggaggaagaataatcgcgaataatgsgtc 840  
Db 781 GATGTCTTTTGCACAGCACACCAAGAGGAGAGAAATGCCAAATATGSGGTC 840

QY 841 ttgttctcaactgagtaacttttgaatttaccgtctatcgcactggccttttggaaaggt 900  
Db 841 TTGTCTTCAACTGAGTACTTTTGGATTTAGCGTGTATCGCACTGGCCTTTTGGAAAGGT 900

QY 901 ttctcagatcatcagaaatggcgaggttgcgcagcgttggcaagtccttactgtctctt 960  
Db 901 TTTGCAATGATATCAGAAATGGCGAGGTTGCCGACGTTGGCAAGTCTTACTGTCTTCCTT 960

QY 961 tccgtcactttagcagccacgtccatctcaatgcttgcgcttcaggttcagtcgctttac 1020  
Db 961 TCCATCCTAATCATGTCTTACGAACAAAAAATCTCTACTTGGTGTCTTCTTGTGTCAACA 1020

Db 961 TCCGTACCTTAGCAGCCACGTCATCTCAATGCTTGCCTTTCAGGTTTCAGTCTTAC 1020

QY 1021 caagccgcacattctcggctccgaattattccagtatacttgacaaacccacgcagctcgac 1080  
Db 1021 CAAGCCGCATCTTGGCTCCGAATTTATTCAGTATATTGACAAACCCACGAGCTCGAC 1080

QY 1081 cctctcgacccctcttgaaagcagccaggggctcctcaggtccaaattgagatcccaaac 1140  
Db 1081 CCTCTGACCCCTTCTGGAAGCAGCAGAGGCTGCCTAGGTCAAAATTGAGATCCAAAAC 1140

QY 1141 ctggcatttgctacccctcccagaccatctgcccgaagtactctcgagatttcaacttgaca 1200  
Db 1141 CTGGCATTTGCTACCTCCCTCCGACCATCTGCCCAAGTACTTCGAGATTTCAACTTGACA 1200

QY 1201 attccagctggcaagcagcagccctcgtcggtagcaggtagcggcaaaagcacaatg 1260  
Db 1201 ATTCAGCTGGCAAGCAGCAGGCTCTGCTGGTGCATCAGGTAGCGGCAAAAGCACAAATG 1260

QY 1261 gtcggcttacttgaaacggtggtatctgcccagttcggggagagatatctactgtgggtg 1320  
Db 1261 GTCGGCTTACTTGAACGGTGGTATCTGCCCAAGTTCTGGGGAGGATATTACTTGTATGGGTTG 1320

QY 1321 gaactgggacaatacaatgtgaaatggctgagaagccgcatctcgcctcgttcaacagga 1380  
Db 1321 GAACTGGGACAATACAATGTGAAATGGCTGAGAAGCCGATTCGGCTCTGTTCAACAGGAA 1380

QY 1381 cctgtgttcttggaacaaatctccagaaatctccagaaatctccagaaatctccagaa 1440  
Db 1381 CCTGTGTGTGTCTGGCACAAATCTTCAGAACATTTGCCAACGGTTTCATGGATGAGCAA 1440

QY 1441 cgagatctgctcgcgaaacaaatggagcttggcaaaagcttggcaaaagcttggcaaaag 1500  
Db 1441 CGAGATCTGCTCCGGAACAAATGGAGCTTGTGCAAAAGCTTGCNAAGCCAGCAAT 1500

QY 1501 ggcgacgtgttcaataatgagcttccgaacggttatgagactgaagtgtggcgagcgacc 1560  
Db 1501 GGGACGCTGTTCATTAATGAGCTTCCGAACGGTTATGAGACTGAAATTTGGCGAGCGACC 1560

QY 1561 ggaacctgagtgagagtcgaacgaacgaattcgaatcgacgaagatcatatcggat 1620  
Db 1561 GGAGCCTTGTAGTGAGGTCAACGACAAGAAATGCAATCGCACGAAGTATCATATCGGAT 1620

QY 1621 cccaagatcctgttactcgtgaagctaccagcgcccttgaccggaagcgcgaggaagtg 1680  
Db 1621 CCCAAGATCTGTGTACTCGATGAAGCTTACCAGCGCCCTTGACCGGAGCGGAGAAAGTG 1680

QY 1681 gtcaggagcccttgaaccgagtgccaaagaccgcactacttctggctcatgtccacaaa 1740  
Db 1681 GTCAGGAGGCTTGAACCGAGTGTCCAAAGACCGCACTACTTTTGGTCAATTTGCCACAAA 1740

QY 1741 ctgacctgtcaaaagtctggcaacatcgacgtcatcttccaggggaaatcgtcgag 1800  
Db 1741 CTAGCCACTGTCAAAAGTGTGGCNACATCGCACTCATTTCCAGGGGAAATCTCGAG 1800

QY 1801 caaggcacacacacgaattgatcgaattcggctgtcattcacgcgcactggtggtgca 1860  
Db 1801 CAAGGCACACACCAAGAAATGATCGAATTCGGCTGTCTTACGCGCACTGCTGCTGCA 1860

QY 1861 caggacctcgggctgacgaacaaagaacatgagagagccctgcacgaaagagcagca 1920  
Db 1861 CAGGACCTCGGGCTGACGAACAAACAAAGAAATGAGAAAGCCCTGACAGGAAAGGACGA 1920

QY 1921 cgagaagctgctggtgaacgacggcacttgagcgactcacacactgcccactctcaaa 1980  
Db 1921 CGAAGAAGTGTGTGAACGACCGGCACTTGAAGGCACTTCAACCACTGCCACATCTCAA 1980

QY 1981 gctggagactggagaagcgaaggtccggtcgggactttgggctactcgtcctcctaaaa 2040  
Db 1981 GCTGGAGACTGGAGAAGCGGAAGGTGCCGCTGGGGACTTTGGGCTACTCTGCTCTCTAAA 2040

QY 2041 tgcactcactatcgttctacgaacaaaaaaactctactcaggtcgtcttcttcttgaaca 2100  
Db 2041 TGCATCCTAATCATGTCTTACGAACAAAAAATCTCTACTTGGTGTCTTCTTGTGTCAACA 2100

```
QY 2101 ataacggtcttgatgcgcgcgcacatttcacagacagcccttbtgttttcgagattg 2160
DB 2101 ATAACGGTCTGATATGCGCGGCCACATTTTCAGGACAAAGCCCTTTTGTTCGAGATTG 2160
QY 2161 ctaactgtcttcgagttgagtggttcacgcgcgcacaggaacggcgagacttttatattcg 2220
DB 2161 CTCACGTGCTTCGAGTTCAGTGGTTCATCGGCACAGGACGGGACAGCTTTTATATCTTG 2220
QY 2221 atgtctcttctgctgctctaggaattctagtagatatttcacagattggctggacatcg 2280
DB 2221 ATGTTCCTTGTGCTGGCTCTAGGAATCTAGTAGGATATTTCCACGATTTGGCTGGACATGC 2280
QY 2281 aacgttattcaacaagttgtcaaccatcgctctatcaagcgcgaatgttccaaacgagtagtg 2340
DB 2281 AACGTTATTTCACAAAGTTGTACCCATCGCTATCAAGCGGCAATGTTCCAAAGAGTACTG 2340
QY 2341 gataagacatgaactcctcagatccgcgagcaaatctctggtctcactcagatcgaa 2400
DB 2341 GATCAAGACATCGAACTCTCGACATCCCGGAGCAAAATTTCTGTGCTCTCACATCGCAA 2400
QY 2401 ctgctcagctctaccacacagtcgaaagttgatatacagcaaatcttctcatttatatc 2460
DB 2401 CTGTCAAGCTCTACCCACGAGTTGCAAGAGTTGATATACGAAATTTTCTCATTTATATC 2460
QY 2461 gttgtcgttcaacatcgctctcagcagtgctctcaccatagcctatggatgaaactggg 2520
DB 2461 GTTGTGCGTCAACATCGTCTCGAGCAGTCTCTACCTAGCTATGGAATGAAACTGGG 2520
QY 2521 cctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2580
DB 2521 CTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY 2581 tctagagcaagctagaaagcgcggaactcgcgaactcttgcaaaactttgcagaagtgctgggtgc 2640
DB 2581 TCTAGAGCAAGCTAGAAAGCGGAACTCGGCAAACTTTGCAAGAAAGTCTGCGGCTGC 2640
QY 2641 aagcgaacagttacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2700
DB 2641 AAGCGAAGCAGTTACCGCGATCGGACCGCTCTCATCTTTGACTCTCGAAGGCCATGTTCT 2700
QY 2701 ccaacgactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2760
DB 2701 CCAACAGTACTCGGACATCTTGAGCAAGCTCTTGCTTAAGATCATCCAAAGCTTTTGTGTT 2760
QY 2761 ggacgatgtttgttttcaactgttcacagtcgagtgagtttttggctatttgcctcgga 2820
DB 2761 GGACGATGTTTGTGTTTCACTTGTACAGTCGATGGAGTGTGCTTATGCTTATGCTCGGA 2820
QY 2821 ttttgtatgcagtcgataattgcttcaggtgagtgagtcacacacactcaattttatca 2880
DB 2821 TTTTGTATTGTCAGTCGATAATTGGCTTCAGGTGAGTACGACACAACTCAATTTTATATCA 2880
QY 2881 tctctgtggcgctttgtttgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2940
DB 2881 TCTTCTGCGGCGTTTGTGCTGCGGTCCAAAGCAGACGCCAGTATTGGCTTACTCCAGC 2940
QY 2941 agttttaccaagctcgttgcgctgcgaactatatacctctggtgcgcgcgcgcgcgcgcgcgc 3000
DB 2941 AGTTTTCACAAAGCTCGGTCGGTGGCACTATATCTCTGCTGCTGGGTGGGACATTTCAAGCCG 3000
QY 3001 accatccgcgaacagagagaaacaaagaaagccagtggtgagtcacctgtcgac 3060
DB 3001 ACCATCCGGAACAGGAGGAGAAACAGAAAAAGGCCAGTGGGTGGATGCCCTGTCGAC 3060
QY 3061 ctgaggaactgaattcaggtatcgtcaacgtgattcgcgtcgagttctccgcggggtt 3120
DB 3061 CTCGAGGACATTAATTCAGGTATCGTCAAGCTGATTGCGCTCGAGTCTCCGCGGGGTT 3120
QY 3121 tccatgacaatcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3180
DB 3121 TCCATGACAATCGACGCAGGACAAATTTGTAGCTTATGTGGCGCTTCTGCTGTGTGGCAAG 3180
```

```
QY 3181 tcaacttcatcgctttgtcggaacgattctcagaccccgacctcgccgacgaatttcattt 3240
DB 3181 TCAACGTTGATCGCTTGTTCGGAACGATTCTACACCCGACCTCGGGCGGCAATTTTCATTT 3240
QY 3241 gcacagagaataatgcagaaatgtcgccgctgttacccgcggccataatgtcttttggtc 3300
DB 3241 GCACAGAGAATAATTCAGAAATGTCGCGCGCTTGTACCGCGGCATATGCTCTTTGCTC 3300
QY 3301 caacagaaacccacacttttccaaagctcgttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3360
DB 3301 CAACAGAAACCCACACTTTTACCAAGGCTTTCGCGGAGATGTGACGTTGGCCCTCGAA 3360
QY 3361 gccgaattatcagaagagctttgtcaagagcgccttccgcgaagcgaatgtctttgatt 3420
DB 3361 GCCGAATTTATCAGAAGAGCTTTGTCAAGAGAGCTTTCGCCGCAAGCCCAATGCTTTGGATT 3420
QY 3421 ttgtcatctcttaccagaaggttttgaacgccttgcgctcgaacgagggatgcagttc 3480
DB 3421 TTGTCTATCTCTTTTACCAGAAAGGCTTTTGAACGCTTTCGCGCTCAACGAGGATGCAGTTC 3480
QY 3481 tccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3540
DB 3481 TCCGCGGCGCAACGACAGCGGATCGCCATCGCAAGAGCATTTGATTCGAAATCCAAAGCTG 3540
QY 3541 ttgtacttgacgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3600
DB 3541 TTGCTACTTCAAGNAGCGAGCTCAGCCCTCGACAGCAATCGGAACGCTCTGTTTCAAGCT 3600
QY 3601 gcctcgatgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3660
DB 3601 GCCTCGATGAGGCATCCAGGAGCGCAACAGCAATAGCAGTGGCGACCGACTTTTCCACT 3660
QY 3661 attcggaattgtgattgtttttgttttgcacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3720
DB 3661 ATTCGGAATGTGATGTTATTTTGTGTTTGTTCACGCGGAGAAATCGCCGAACGGGCACT 3720
QY 3721 caqcggaactacacgactgagaggaagatattacagatggttttggcacaactcttta 3780
DB 3721 CACCGGAACACTAACGACTGAGAGGAAGATATTACGAGATGTTGTTGGCACAAATCTTTA 3780
QY 3781 gaccaagcatga 3792
DB 3781 GACCAAGCATGA 3792
```

## RESULT 2

```
AX076850
LOCUS AX076850 3999 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 9 from Patent WO0105980.
ACCESSION AX076850
VERSION AX076850.1 GI:13121522
KEYWORDS
SOURCE Exophiala spinifera.
ORGANISM Exophiala spinifera
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothyriomycetes;
Chaetothyriales; Herpotrichiellaceae; mitosporic
Herpotrichiellaceae; Exophiala.
REFERENCE 1 (bases 1 to 3999)
AUTHORS Duvick,J.P., Maddox,J., Gilliam,J., Folkers,O. and Crasta,O.R.
TITLE Compositions and methods for fumonisin detoxification
JOURNAL Patent: WO 0105980-A 9 25-JAN-2001;
Pioneer Hi-Bred International Inc. (US) ; Curagen Corporation (US)
FEATURES
source 1..3999
/organism="Exophiala spinifera"
/db_xref="taxon:91928"
/note="p-glycoprotein, with introns"
BASE COUNT 1007 a 1008 c 992 g 987 t 5 others
ORIGIN
```

Query Match 91.1%; Score 3454.8; DB 6; Length 3999;  
Best Local Similarity 96.1%; Pred. No. 0;



QY	2150	tttcgagattgctcactgtcttcgagtttgagtggtctatcgcgacagaaaggcgacact	2200
Db	2178	TTTCGAGAAATGCTACTGCTTTCGAGTTTGAAGTGGTCATCGCGCACAGGAACGGGCAGACT	2237
QY	2210	tttatattctgatgtctcttcttgctcggtgctctaggaaatcttagtaggatatattcaagattg	2269
Db	2238	TTTATAGTCTGATGTTTCTTGTGTCGTGGCTTAGGAAATCTAGTAGGATATTTTCAAGATTG	2297
QY	2270	gctggacatgcgaacggttatttcacaagtttgccaccatcgctatcaagcgcgaatggtcc	2329
Db	2298	GCTGGACATGCAACGTTGTTTTCACAAGTTGTACACCTATCGCTATCGAGCGGAATGTTCC	2357
QY	2330	aacgagtactggatcaagacatcgaactcctcgcagatccccggagcaacaatttctgggtctc	2389
Db	2358	AACGAGTACTGGATCAAGACATCGAAATTTTCGCATATCCCGGAAATACTTCTTGGTGCTC	2417
QY	2390	tcacatcgcaactgtcagctctaccacgcagcttgcaagagttgatatacagcaaat-tt	2448
Db	2418	TCACATCGCAACTGTGAGCTCTACCCACGCGAGTTGCAGGAGTTGATATCAACAAATTTCTT	2477
QY	2449	ctcatittatcgttgtcgggtcaacatcgtctcgagaagtgctctcacactagactatggtg	2508
Db	2478	CTCATTTTATCGTTGCTGTACAACATCCTCTCGAGCAGTGTCTTAGCATTAGCATTAGG	2537
QY	2509	atggaaactgggcctgggtgtgttgggtgcacttcacccctgctcttggctggctcta	2568
Db	2538	ATGGAACATGGGCGCTGGTGGTTGTTGGTGCATCTCCACCCCTGCTTTGGCTGGCTA	2597
QY	2569	cctcagaattcgtctagagacgaagctagaagcgggaactcggcaaaacttgcagaag	2628
Db	2598	CCTCAGAATTCGCTTGTAGACGAAGCTAGAAGCCGGAACCTCGGCAAACTTTGCAGAAAG	2657
QY	2629	tgctgggcttgcaagcgaagcagtttaccgcgcatccgacgctctcatcttgactctcga	2688
Db	2658	TGCTGGGCTTGCACCGAAGACACTTACCCCGATCCGGACCCGTCACACTTGGACTCTCGA	2717
QY	2689	aggccaattgcttcacaagctactcggacatgttggaagaaggtcttgtaagaatcattgcata	2748
Db	2718	AGGSCATGTCCTCCAAACAGTACTCGGACATGTTGAGCAAGGTCTGCTGAAGATCATCCAA	2777
QY	2749	agctttgtttggacgagttgtttggttttccacttgcacagtcgatggagttttggttggct	2808
Db	2778	AGC-TTTGGTTTGGACGATGTTTGGTTCTCAC-TGTCACAGTCGATCGAGTTTCTGGCT	2835
QY	2809	attgcccctgggatttttgat---tgacgtcgataattggcttcagggtga-gtacgacaca	2864
Db	2836	ATGGCCCTGGGAATTTTGGTATGGGAAGTCGACACTTACGGCTTCAGGTGAGGTACGACACA	2895
QY	2865	actcaatttatatcatctctgtgggcgttttggtttgcgggtccaaagcagcagcccaagta	2924
Db	2896	ACTCAATTTTATATCATCTGCTGGGCGTGTGTTTGGCTGCCGTCCAAAGCAGACGCCAGAA	2955
QY	2925	tttggcttactccacgagttttacaaaggctcgggtcgagctgcgaactatctctctggct	2984
Db	2956	GCCGAATTACTCCACGAGCTTTACCAAGCTCGGTGGCTGCCGAATATATCTCTGGCT	3015
QY	2985	gcggacattgaagccgaccatccgcgaacccgaggaacaagaataaaagcccgagtggg	3044
Db	3016	GCGGACATTTGAAGCCGACCATCCCGGAACCGGAGAGAACAAAGAAAAAAGGGCGCAGTGGG	3075
QY	3045	tggatgccctgtcgacctcgaggacattgaattcaggttatcgtcaacggtattcgctcg	3104
Db	3076	TGGATGCCCTGTGCGACCTCGAGGACATTTGAATTCAGGTATCGTCAACGTGATTCGGGCTCG	3135
QY	3105	agttctccgcggggtttccatgacaatcgagccaggacaatttgttagcttattgtgggcgc	3164
Db	3136	AGTTCCTCCGGGGGTTTCCATGACAAATCGAGCCAGGACAAATTTGTAGCTTATGTGGGCGC	3195
QY	3165	tctcgtgtgcgaagtcgaagttgatcgttttgcggaaagattctaacgacccgacctc	3224
Db	3196	TTCTGCTGTGCAGAGTCAACGTTGATCGCTTTGTTGGAAGGATTTTACGACCCGACCTC	3255
QY	3225	gggcccgaatttcatttgcacacgagaattatgcagaagtgcgcgcgcgttgcacgcgcg	3284

[illegible]

```
/product="ATP-binding cassette multidrug transport protein
ATRC"
<628. .>4704
/gene="atrc"
CDS
join(628..747,813..1000,1047..1703,1761..3948,4003..4704)
/gene="atrc"
/function="putatively involved in multidrug transport"
/notes="similar to MDR; ABC transporter C"
/codon_start=1
/product="ATP-binding cassette multidrug transport protein
ATRC"
/protein_id="AAD43625.1"
/db_xref="GI:5456699"
/translation="MKSTAESKETPSODESTTVPCTEAPLVEGEASRGAVKRIFF
FAGRLIIQAVAILAACAGAGIALQNLIFGQFVTVITDFTNGIISPADFRONAEEL
ALFVYVGLKIQAGSDKICLLFQGLAFTLSLRUWCKWKLFLICIPVATIGTG
VYNAEPAGHETRIQIHAQANSFAEGLIAGVAVHAFGRMDLSVRDFEVLBAHVGL
KKISPLILGLFSAEYTIYILGYGLAFWQGIHMFGEIGTAGDIFTVLISVLAISNL
TLIAPYSIEFSAASAAQIIFRLIDRESEINPXKRGLEPERVLGVELENVTFSTPT
RPGITVDNFSLSKPAKVTLVQSGSGKSTIVGLLEWYNTPTSGAIBLDGNLISEL
NVGLRRNVRLVQEPVLFQGSVFQNIYVGLVTPHENASREOMERVQAKLATAH
EFISELTDGYDLTILGERGLLGGQKQVIAIRSVVQPKVLLDDEATSAIDPHAETI
VOKALDKAAEGRITTVIAHKLATIRKADNIVMSKGHIVEQSTHSLIAKQDVIYAGLV
KIQNLAVNASADHNNEGEDEVALLEVTEATVTRYPTSIRGRMNSIKRDDYENHK
HMDMLAAALAYLRECEPELKWAYLVLLGLGCGCAMPYGOAILMSRVVEVTSGDAML
DKGDFYASMLIVLAAGCLICYLAVGATNTIAQHLSHFRRLIHLMDLRQDIOFDDRE
ENTTCALVSRIYSPHAILLMGYNIALVAVLVVTCGILAIAPFSWKLGLVYVFGG
IPLVAGVAVRVRSDRDROTSKKYGTSSTASEAVNAIRVSSLAISETVLRRYTE
ELDHAVSSVKPMAATHMICFGLTQCIWFQALGFWIGKRLVSLGTSWYFFVFLS
VFQAGSAQLFOWSTSTKGINATNYIAWLHQLQPTVRETPEHNDKFGSGAPIAMD
NVRFSYPLRPDAPILKGNVNLKNGOFIAFVSGSGCKSTMIAMLERFVDPDTGSI
DASTLTDNRINRYNVALVOEPIYTFQGTIRDNISLVFNPTOPFFSDKDAKSVS
DEQIESALRANAFVSLPOGIYTPAGSGGSLSGQORQIATARALIRPKILL
DEATSAIDTESKIVOKALEAGDRDLTVAVHRLSTIKDANVICVFFEGKIAEMG
THQELIVRGLLYRMCQAQALD"
BASE COUNT      1497 a 1531 c 1608 g 1484 t
ORIGIN
```

```
Query Match      14.3%; Score 543.8; DB 8; Length 6120;
Best Local Similarity 50.5%; Pred. No. 2,3e-141;
Matches 1854; Conservative 1; Mismatches 1643; Indels 175; Gaps 15;

QY 289 ttcgtctggtggttcacgttattgggaagttgttcctcagtcagctttccaag 348
DB 1042 TTCAGCCCTCTACTTGTATACCTGGGCATCGCGCGCTCGTCTCTCTACACCTACAAC 1101

QY 349 gctgcattaccatttcagctatcagaaccactcgaaactcttcgaagagtggttccttga 408
DB 1102 ACCCTCTTAACCTACGCGGCCCTACCGCATCTGTCGCGCAATATCCGACACGCCCTATCTCAA 1161

QY 409 tgcaccttggcgaagaggtctggcatttcgacaagcagacgaatggagcaatcgccact 468
DB 1162 GCGGCGCTGAGCAAGAGTGGCATCTACGATTTTCGGTAGCGGGGGCTCCATCGCGCGG 1221

QY 469 cargtcaactacaatggcaacogtatacaaacaggattatggcgaagaattgcttcttacc 528
DB 1222 CAGGCAACTTCGAACGGCAAACTGTATCCAGCGCGCGCTCGGATGAAGATCGTCTCTCTC 1281

QY 529 gtgcaggcacttctcaattcttcctgatttggttgcgttttggtcgtctcagtggaag 588
DB 1282 TTCAGGCGCTCGGAGCAATGTGTGACGCTTTCATTATCGCGTTTGTGTGCAAGTGGAA 1341

QY 589 ctagctttaatcacatgctcgcgtatccctgccatttctcgttggtcacggcaatcgcata 648
DB 1342 CTCACCTCTGATCTGCATCTGATCCCCGTAGCCACGATCGGCACACGGGGGTGTAGCT 1401

QY 649 gcaattgatcgctcaggagcaggaatcaccaggatctactcagcgcgctgctctc 708
DB 1402 CGGTCGAGGCTGGCGACGACGAGGATCTTGAGATACATGCGCAGCGCGGAATTCGTT 1461

QY 709 gcagaagaagcttatcatccatccggacagtcctatgcttctcagcccgcaagaagaatg 768
```

```
DB 1462 GCCGAGGGTATTTCTGGCGGGTGTGAAGGCTGTTTCATGCTTTTGGGATCGGGATAGTCTG 1521
QY 769 gtgcgaataatgatgtctctttttgcagcagcacacacaagaagggaagaagaatcgcca 828
DB 1522 GTCAGGAAGTTTGTATGAATATCTGTTGGAGCGCATAGGTCGTAAGAAGATCTCGCCG 1581
QY 829 aataatgsgtctcttcttcaactgagtgacttttggaattgaactgactgctatcgactggcc 888
DB 1582 CTGCTTGTGCTCTCTTCTTCGCGGGAGTATACGATCATCTACCTTGGATATGGCTGGCG 1641
QY 889 ttttggaaaggttttcgcatgtatcagaatgcgaggtttgcgcagcttggcaaaagctttt 948
DB 1642 TTTTGGCAGGGATCCATATGTTCCGCAGGGGGAGATTGGGACTCTCTGGGATATCTTT 1701
QY 949 act----- 951
DB 1702 ACGTACTACGACCCCTTTCCCTTTTTCAGCGATCAGTGAAGATATGACATGACTCAACAGG 1761
QY 952 gttgccttcttcgtaacctttagcagcagcgcgtccatctcaatcttgcgcttcaggttca 1011
DB 1762 GTTTTGTCTCTGTGCTCATTTGCGTCAATCAACCTGACTTTTACTGCGCGCTATTTCATTT 1821
QY 1012 gtcgtttaccacaagcgcgcattcttcggtccg-aattatcagtcatactgacaaacccac 1070
DB 1822 GA-ATTTAGCAGGGTCTGTTTCAGCGGCTGCGCAACTGTTCCGACTCATAGATCGAGAGTC 1880
QY 1071 gcactgcgacctctgcacctcttgcgaagcagcaggggctgcctaggttcaaatga 1130
DB 1881 TGAATCAACCCATACCGGAAGAGGCTCTGAGCGCGAAGGGTATTAGCGCAGCTCGA 1940
QY 1131 gatccaaacctgycatttgcctacctcccgacacatctgccaaagtacttcagagattt 1190
DB 1941 GCTCGAAGTGTACGTCTCTGATCCACGAGGCGCGGGATTACCGTCTCTCGATAAATT 2000
QY 1191 caacttgacaattccagctggcgaagcagcgcgctcgtcgtgcatcaggttagcggca 1250
DB 2001 CAGTCTCAAGTCCAGCGGGAAGGTGACTGCCCTGTAGGCAATCTGGATCGGGAA 2060
QY 1251 aagcacaatggtcggttacttgaacgggtggttatctgcacagtttcggggaagataattact 1310
DB 2061 GAGCACATCGTGGGATTTGCTCGAGCGGTGTTATAACCCGACCTCTTGGGGGATCAGACT 2120
QY 1311 tgatgggttgaaactgggacaatacaatgtgaatggctgagaaagcgcattcgcctcgt 1370
DB 2121 CGAGGGAACCTGTATGAGTCAATGTTGGCTGGCTCGGAGGAATGTGCGGCTCGT 2180
QY 1371 tcaacaggaacctgtgttcttgcgttcgacaaatcttccagaaacattgccaacggtttcat 1430
DB 2181 ACAGCAGAGCGCGTCTTCCAGGGAAGGCTGTTGATTAACATCAGTACGGCTCGT 2240
QY 1431 gtagtgaacagag---atctgcctcgcgaaaaaaataatggagcttgcgaaaaagcttg 1487
DB 2241 CGGAGCGCTCGGAGAAATGCCCTCTCGGGAAGACAGATGGAAGGGGTGAGAGGCGCG 2300
QY 1488 caaagcagcaatggcgacgtgttcattaatgagcttccgacggttatgagactgaagt 1547
DB 2301 GAAGTTGGCATATGCGCACGCAATTCATCTGAGCTGACCGACGGATACGATACGCTGAT 2360
QY 1548 tggcgagcagcgggagccttgcgtgaggttcaacgcgaacgaattgcaatcgacacgaag 1607
DB 2361 CGCGAAGCGGGTGTCTGCTTCTGAGGCCAGAACGAGCAGCGGGTTCGGAATTCGCCGCG 2420
QY 1608 tatcatatcggaatcccaagatcctgttactcagatgaagctaccagcgcccttgaccgaa 1667
DB 2421 CGTCGTTTCTCAACCGAAGGTCTCTTCTGCTGGATGAAGCAACACAGTCTCTTGTATCGCA 2480
QY 1668 ggcggagaagtggtccagagggccttgaaaccgagtgctccaaagacgcgactacttgg 1727
DB 2481 TGCAGAGCATGTTTTCAGAAAGCTCTGGACAAAGCAGCTGAGGGGCGCAGCAGCATGTG 2540
QY 1728 cattgcccaacaaactgacactgtcgaagtgtcgaacatcgacagtcatttccccaggg 1787
```



Db	2541	CATTGCTCAAAACTTTGCTACGATCCGAAGCGGACAAATATCGTTGTCATGACAAAGG	2600
Qy	1788	gaaatcgtcgagcaaggcacacacacgaattgatcgaaattcgctgtcattacgcgcg	1847
Db	2601	TCACATTGTCCGAAGGCACACAGAGTCACTGATACCCAAAGACGCGCTCTATGCGCG	2660
Qy	1848	actgggtgctgcacaggaacctcgggctgacgaacacacagacatgagaaagacctgcga	1907
Db	2661	TCGTGGTCAAAATCCGAACCTGGCAGTGAATGCTTTACACACATGACAAATGTAATAGSAG	2720
Qy	1908	cgaaaaggcagcagcagaaagt-----gctgggtgaacgcagccgcacctgagcgcactca	1961
Db	2721	GGGTGAAGGCGAAGATGTCGCTCTCCTGGAGGTCACCGAAACAGCAGTAAACCGCTACCC	2780
Qy	1962	cacacgtgccacatctcaagctggaagacctggagacggaagtcgctgcggacctt	2021
Db	2781	AACTCCATCCGCGGTTCGAATGAACATCCATAAAGACCGCAGCATATATGAGAACCAAA	2840
Qy	2022	gggctactcgctctaaatgcataatcatgcttctacgaacaaataatctctactg	2081
Db	2841	GCACATGGGATATGCTGGCGCCCTTAGCTTAATCTCGTCGCGCGAATGTCGAGACTGGAATG	2900
Qy	2082	gtgcttcttgtgtcaacaataaacggttctgatatgcgggccacatttccaggacaagc	2141
Db	2901	GGCCTATCTCGTCTGCTACTGGGTGTCTTGGTGTGGCCATGACCCCGGCCAAGC	2960
Qy	2142	cccttgttttcgagaattgctcaactgttcttcagattgagttgctatgcgcgcagaagacg	2201
Db	2961	TATCTTGAATGTCFCGGTGTTCGAGGTCTTCAGCGTCFCGGGACACGCTATGCTAGACAA	3020
Qy	2202	ggcagacttttatactctgatgttcttctgtcgtctaggaatactagtaggatatatt	2261
Db	3021	AGGAGACTTCTATGCCAGTATGCTGATCGTCTCGGCGCCGGGTGCTGATCTGTACTT	3080
Qy	2262	cacgattgctctggagacgcgaagcttatttcaaaagtgttaccacatcgctatcaagcgcg	2321
Db	3081	AGCTGTCCGATATGCAACCAACACACTATAGCCAGCATTTAGTCAATTGGTTTCGACGCC	3140
Qy	2322	aatgttccaacagtgactggatacgaagacatcgaaactctcgacatccgcgagcaaattc	2381
Db	3141	CAITTCGCAGCATGCTCGGACAGATACAGTTCCTTGACCGTGAAGAGAACACTAC	3200
Qy	2382	tgtgtctccacatcgcaactgctagctctaccacgcagttgcaagagttgatatacgc	2441
Db	3201	CGGTGCGCTGGTAAAGCGTATCGATTCGTACCGCATGCAATTCGAACTTGGGCTA	3260
Qy	2442	aaatttccatttatatcgttgtcgctcaaca--tcgtctcgagcagtgctctaccaat	2499
Db	3261	CAACATCGCCCTGGTGTGATTTGCTGCTCCGAGGTGGTAACCTGTGGCATTCCTTGGCCAT	3320
Qy	2500	agcctatgatggaaactggcctgggtgtgtgtgttgatgcattccacccccgtttt	2559
Db	3321	TGCATTCTCCTGGAAACTAAGGCGTGGTGTCTTTTGGCGGTATTCCACCCCTTGTTCGG	3380
Qy	2560	ggctggtctaccloagaattcgtctagagacgaagctagaacgcggaaactcgcaaaact	2619
Db	3381	TGCTGGATGCTAGCAATCCGCGTCACTCCGCGCTCGATCGCCAGACATCGAAGAAATA	3440
Qy	2620	tgcagaaagtgtcgtggccttgcagcgaagcaggttaccgcgcatccggaccgtctcatcttt	2679
Db	3441	TGGCACCAGCTCTCCATTGCCCTCTGAAGCTGAACGCTATCCGGACCGTTTTCGTCCCT	3500
Qy	2680	gactctcgaagcccatgttctccacagtaactcggacatgttgagcaaggtcttctgtaag	2739
Db	3501	TGCAATCGAAGAGACGGTCTACGTCGATACACGGAGGAACTAGACACGCTGCTGTCTCGT	3559
Qy	2740	atcatccaaagcttttgttgagcagatgttttgggttttcaacttgctcacagtcgagtgag	2799
Db	3560	CTTCGSGTGAACCCATGGCTGCCAGATGATTT--GTTTCGGGCTGACCGAGTCGATTGAG	3618
Qy	2800	tttttggctatgcccctgggatt-ttgtattgcagtcgatataattggcttcaggtaggtac	2858
Db	3619	TACTGTTTTACGCGCTGGGATTTCTGGTATGGGTGTCTGCTTGTGTGCTGCTGGGGAGACT	3678

QY	2859	gacacaactcaatttatatcatctctcgtggtggcggttttgtttccgggtcccaagcagcgc	2918
Db	3679	ACCATATATAGTTTCTTTGTTCGATCTCCTCAGATGTGTCTTTTCCGGGT-CAGGCGTCA	3737
QY	2919	ccagttattggcttactccaagagttttccaaggtccggtcggtcgcaactatacct	2978
Db	3738	CGAGCTGTTCCAGTGTTCGACCAAGTATTACAAAGGGAATCAATCGGACGAATACATCG	3797
QY	2979	ctggctcgcggaacattgaagccgaccattccgcgaacgagagagaacaagaaaaagggccc	3038
Db	3798	TTGGTTGTCACACAGCTCCAACCAACAGTGCOCGAGACGCGGGAACACCATTAAGGCC	3857
QY	3039	agtgggtggatgcctcgtcgacctcgaggacattgaattcaggtatcgtccaagctgattc	3098
Db	3858	TGGATCTGGGGGCCGATGCTATATGACAAATGTGCGCTTCTCGTACCTCTACGGCCAGA	3917
QY	3099	ggctcgaagtctccgcggggttccatgaca-----	3129
Db	3918	CGCCCCATCTCCTGAAAGGGGTGAATCTGAAGGTAAACAATCATCCCGGAATATGATCAAT	3977
QY	3130	-----atcgagccagagacaattgtagcttatgtggcgcg	3164
Db	3978	CAAAAGATACTAATCGATACGAACAGATAAAACAAGGCCAATTCATCGCTTCGTCGGCTC	4037
QY	3165	ttctgctctggcaagtcgaactgacgtctgttcggaacgattctcaagccgacccctc	3224
Db	4038	CTCCGGCTCGCGCAATCACCACATGATGTCATGCTCGAGCGCTTCTACGATCCCAACAC	4097
QY	3225	ggcgcaaatctcatttgcacacgagaaatttcgagaaatgtcgccgctctgaccgcgg	3284
Db	4098	AGGGAGCATACAAATCGACGCTTCCACCTCACCGACATAAACCCCATATCTTACCGAAA	4157
QY	3285	ccatatgtcttggtccaacaggaacccacactttcccaaggtctcgttcgcgagaaatgt	3344
Db	4158	TATGTGGCAGCTGGTCAGCAGAGGCCAACCCCTTTTCCAAAGGGACAAATACGGGACAACAT	4217
QY	3345	gacgttggccctcg--aagccgaattatcagaagagcttt-----	3382
Db	4218	CTCGTGTGGCTCTTCAACCCGGAATACGACGCCCTTTCTTTCTGATAAAGGATCGAGTAA	4277
QY	3383	-----gtcaagagacgctctccgcaagggcaaatgcttggatttgt	3424
Db	4278	GTCCGTGCTGTATGACGAGATTCAGTCGGCCCTCCCGCGAGCTTAATGCCCTGGGACCTTGT	4337
QY	3425	catctcttaccagaaggctttgaaacgcttgcggtccaacgaggagtcagttctcg	3484
Db	4338	CTCTCTATTTCGCGCAGGGATCTACACGCCGCCGTGGCTC-AGCGGGGTCCCAACTCTCTG	4396
QY	3485	gcgggcaacgacagcgtatgcctatgcgaagacattgattcgaatccaaagctattgc	3544
Db	4397	GGGGCAGCGGCNACGCANTGCCATTTGCCCGCGCTCATCCGAGATCCAAAGATCTTAC	4456
QY	3545	tacttgacgaagcgcgtcagccctgcacgcgcaatcggaacgtctggttccaagctgccc	3604
Db	4457	TCTTTGACGAGGCTACGAGTCTGGATACAGAGAGTGAGAAAGATCGTGACAAAGGCTC	4516
QY	3605	tcgatgagcattccacgag-----ccgaacgacaatagcagtggcgacccgacttcca	3658
Db	4517	TCGAGGGCGCGCCAGGACGGGGACCGGCTTACGGTTGCTGTTCGCATCGATTAAGCA	4576
QY	3659	ctattcggaattgtgatttattttgtttgtccaacgggagaatcgccgaacgggca	3718
Db	4577	CGATTAAAGATGCTAATGTTATCTGTGTATTCTTTTCGGAGGAAAGATTTGGGAGATGGAA	4636
QY	3719	ctcacgcggaactacaacgactgagaggaagattaccagatgtgttggcacaactct	3778
Db	4637	CGCATCAAGAGTTAATAGTTAGGGGGGGCTGTATAGACGGATGTGTGAGGCCACGCCCT	4696
QY	3779	tagaccaagcatg	3791
Db	4697	TGGACTAAGAATG	4709

RESULT 4  
AR149776  
LOCUS AR149776 4002 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 1 from patent US 6228615.  
ACCESSION AR149776  
VERSION AR149776.1 GI:15114367  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 4002)  
AUTHORS Skatrud,P.Luther, de Waard,M.A., Andrade,A.C. and Peery,R.Brown.  
TITLE Multiple drug resistance gene atrD of Aspergillus nidulans  
JOURNAL Patent: US 6228615-A 1 08-MAY-2001;  
FEATURES  
Location/Qualifiers  
1..4002  
/organism="unknown"  
BASE COUNT 974 a 1034 c 1007 g 987 t  
ORIGIN  
Query Match 12.8%; Score 487; DB 6; Length 4002;  
Best local similarity 48.6%; Pred.No.2.1e-125;  
Matches 1731; Conservative 0; Mismatches 1760; Indels 68; Gaps 12;  
QY 262 gaagcttcgaagccgatgttgatcatcttcgtctggttgctctacacctttattggg 321  
DB 424 GACGAGTTCTATGATGAATGACCAAGACGTACTGTACTTCTGTATACCTCGGTATCGGC 483  
QY 322 aagtttgtctcagtcagtttccagcgtgcattaccattccattcagatatcagaaccact 381  
DB 484 GAGTTTGTCACTGCTATGTGTAGTCTGTGTGGTTCATCTATACCGGAGAACACGCCACG 543  
QY 382 cgaactcttcgacgagttgtcttgatgcaccttgcggcaagaggtctgtgcatcttcgac 441  
DB 544 CAGAGATCCGCGAGATTACCTTTGAGTCTATCTCTGGCCGAGACATTTGGCTATTTGAT 603  
QY 442 aagcagagcaatgagcaatcgcaatcargtcactaccacaaatggcaaccgtatcacaaaca 501  
DB 604 AAACTCGGTCCGGGGAAGTAGCACCCCGTATAACAGCCGATACAAACCTTATCCAGGAT 663  
QY 502 ggtattgcgagaaaattggtctttaccgtgcagggcactttcaatgctttcttttgcattt 561  
DB 664 GCATTTCCGAGAAAGTCGGTCTCACTTTGACTGGCCCTGGCGACATTCGTCGACGATTC 723  
QY 562 gtggtgtgttggtctcagtggaagctagcttttaaccacatgtccgtcatccctgcc 621  
DB 724 ATTATGCCCTACGTCAATACCTGGAAGTTGGCTCTAATTTCAGCTCAACAATTGGGCC 783  
QY 622 attttcctgttcaccggcatctgcataagcaattgactgcgtcagagagccaggatcaacc 681  
DB 784 CTCGTCTCACCATGGGGCGGTGGTCTCAGTTTATCATCAAGTACACGAAAGTGGCTT 843  
QY 682 aggatctactcacgcgcgcgtctccctcgcagagaagctcttatccatccatccggacagtc 741  
DB 844 GACAGCTACGGTGCAGGGGCACTGTTGGGGAAGAGGTCTACAGCTCCATCAGAAATGCC 903  
QY 742 catgtttctacgccagaagaaaatggtcgaataatgatgtcttttttcagcagaagca 801  
DB 904 ACAGCGTTTGGCACCCCAAGACAAGCTTGGGAAGCAGTATGAGGTCCACTTAGACGAAGCT 963  
QY 802 caccagaaggaagaaaatcgcaataatgsgtctgtttctcaactagtaacttt 861  
DB 964 GAGAAATGGGNAACAAGAACAGATTGTCAATGGGTTCATGATTTGGCCCATTTGGCC 1023  
QY 862 tgcattacgtgtatgcactgacctgttggaaaggttttcgcagtgtatcagaatgac 921  
DB 1024 CTTATGTACTCGAACTACGGTCTTGGCTTCGTGATGGGTCTCTCGTTCTCTGTAGATGT 1083  
QY 922 gaggttgcagcgttggaagcttttactgttgccttccctcactcacttagcagcaag 981  
DB 1084 GCAGT---CGATGTGGGTGATATTCTACAGTTCTCTCATGGCCATCTTGATCGGATCGTTTC 1140

QY 982 tccatctcaatgcttgccgcttcagggttcagtgctgtttaccacgcgcgcacatcttctgggtcc 1041  
DB 1141 TCTTGGGGAACGTTAGTCCAAATGCTCAAGCATTTTACAACACGCTGTGGCGCGCGCGCA 1200  
QY 1042 gaattattcagtatcatgacaaaacccacgcagcagctctcgacctctcgacctctctggaaag 1101  
DB 1201 AGATATTGGAAACGATCGATCGCCAGTCCCATTTAGATCCATATTGGAACGAAGGGAAG 1260  
QY 1102 cagccagaggggtgcttaggttcaaatggatccaaaacccctggcatttgcctacccctcc 1161  
DB 1261 AGCGCTCGACCATTTTGGGGCCACATTTAGTTACGCAATGCTCAAGCATATTTACCCATCT 1320  
QY 1162 cgaccatctgcccagtaocttcagagatttcaacttgacaattccacagctggcagaagcag 1221  
DB 1321 AGCCCGAGGTCAACCGTCAATGGAGGATGTTCTCTGTCATATGCCCTTGGAAAAACAAC 1380  
QY 1222 gccctcgtcgggtgcatcaggtagcggcaaaagcacaatggtcggtcttacttgaaacggtg 1281  
DB 1381 GCTTTAGTCGGCCCTCTGGCTCTGAAAAAGTACGGTGGTGGCTTGGTTGAGCGATTTC 1440  
QY 1282 tatctgcccagctcggggagagatatctattgatgggttggaactgggacaatacaatgtg 1341  
DB 1441 TACATGCCCTGTTCGGGTACGGTTTTTGTGTGATGGCATGACATCAAGGACCTCAATCTC 1500  
QY 1342 aaatggctggaagccgccttcgccttcacacaggaacctgtgtgttttctgttggtgacaca 1401  
DB 1501 CGCTGGCTTCGCCACACAGATCTCTTTGGTTACCCAGAGCCGTGCTTTTGGACGACG 1560  
QY 1402 atcttcagaacatgcccacgggtttcatgg---atgagcaacagagatcgtcctcgcga 1458  
DB 1561 ATTTATAAGATATTAGGCAAGGTCTCATCGGCACAAAGTACGAGAAATGAATCCGAGGAT 1620  
QY 1459 aaacaatggagctgtgcaaaaagcttgcaaaagccagcaatggcgagcgtgttcattat 1518  
DB 1621 AAGTCCGGGAACATCATCGAAGACGCGCAAAATGGCGAATGCTCATGACTTTATTACT 1680  
QY 1519 gagcttcgaacgggttatgagactgaagttggtcgagcagcgagcggttgggtgaggt 1578  
DB 1681 GCCTTGCTGAAGGTATGACACCAATCTTGGGACGCTGGCTTCTCTTTCAGGTGGC 1740  
QY 1579 caacgacaacgaatgcaatcgacgaagtatcatatcggtatcccaagatcctgttactc 1638  
DB 1741 CAGAAACAGCGCATTTGCAATCGCCGCTGCTGTTGTAGTAGCCCAAAAATCCTGCTCTG 1800  
QY 1639 gatgaagctaccagcgccttgaccgaagcgggagaagtggtccagagagccttgaac 1698  
DB 1801 GATGAAGCTACTTCGGCCTTGGACACAAAATCCGAAGCGGTGGTTCAAGCAGCTTTGGAG 1860  
QY 1699 cgagttccaaaagaccgcactactttggtcatgtcccaacaaactagccactgtcaaaagt 1758  
DB 1861 AGGCAAGCTGAAGCGCAACTACTATTGTGATCGCTCATCGCTTTCCACGATCAAAAGC 1920  
QY 1759 gctggcaacatcgacgtcatttcccgaggggaaaaatcgctcgagcagagcacaacacacgaa 1818  
DB 1921 GCGCACAAATTTGTTCTTGGTCAATGGCAAAATGCTGAACAAGGAACACACGATGAA 1980  
QY 1819 ttgatgaattcggctgtcattacgcgcgcactggtggtgac----- 1861  
DB 1981 TTGTTGACCGCGGAGCGCTTATCGCANAACTTGTGGNGGCTCAACGTATCAATGAACAG 2040  
QY 1862 -----aggacctcgggtggtgacgaacacaagaacatgagaagacc 1902  
DB 2041 AAGAAAGCTGACGCCCTTGGAGGAGCGCGCTGAGGATCTCAGGAATGCAAGATATTGCC 2100  
QY 1903 ctgcagaaaagcagcagcagagctgctggtgaaagccgcacgttgagcgcactcac 1962  
DB 2101 AAATCAAACTCGGTCAACCGCATCATCGATCTCGACGAAACCCCAACCACTTGAC 2160  
QY 1963 accactgccacatctcaagctggagacgttgagaagc-----ggaaggtgctcgggtcgg 2016  
DB 2161 CGCACGGGACCCACAAGTCTGTTTCCAGCGCGATTCTTCTTAAAGAGACCCCGCAACA 2220



QY 262 gaagcggtcaagccgagatgagatcatttcgtctctgtggttgctacacctctttattggg 321  
Db 424 GACGAGTTCTATGATGAATACCAAGAACGACTACTACTCTGATATACCTCGGTATCGGC 483  
QY 322 aagtttgctcagctacgttttccacggctgccattaccatttcagctatcacgaacact 381  
Db 484 GAGTTTGTCACTGTCTATGTTTACTGTGTGGCTTCATCTATACCGGAGAACACGCCACG 543  
QY 382 cgaactcttcgacgagtgctcttgaatgcactcttgccgaagagtgctggcacttcogac 441  
Db 544 CAGAAGATCCGGGAGTATTACCTTGAGTCTATCTCTGCGCCAGACATATGGCTATTTTGAT 603  
QY 442 aagcagagcaatgagcaatcgccactcargtcaactaccatggaacacggtatcacaaaca 501  
Db 604 AAACCTGGTGGCGGGAGTAGTACCCCGTATTAACAGCCGATACAAACTTATCCAGGAT 663  
QY 502 ggtatcgcaagaattggtcttaccgtgcaaggaactttcaatggttcttttctgcatitt 561  
Db 664 GGCATTTCCGGAAGGTGGTCTACATTTTGACTGGCTGGCGACATTTCTGTGACAGCATTC 723  
QY 562 gtggctgctttggctcagtggaagctagctttaaaccacatgctccgtcatccctgcc 621  
Db 724 ATTATCGCCTACGTCNAATACTGGAAGTTGGCTCTAATTTTGAGCTCAACAATTTGGCC 783  
QY 622 atttctcgttaccgggaatctgcatagcaattgatgcgctcagaggagccaggaatcac 681  
Db 784 CTCGTCTCACCATGGCGGTGGTCTCAGTTTATCATCAAGTACAGCAAAAAGTCGCTT 843  
QY 682 aggtactactcagcgcgctgtctctcgcagaagaagttcttatcatccatccggacagtc 741  
Db 844 GACAGCTACGGTGCAGGGCGGCACTGTTCGGGAAGAGGTCTACAGCTCCATCAAGAAATGCC 903  
QY 742 catgctttctacgcccaagaagaaatggtcgaaaaatgatgtctcttttgcagaagca 801  
Db 904 ACAGCGTTTGGCACCCACAGACAGCTTGGGAAGCAGTATGAGGTCCACTTAGACGAGCT 963  
QY 802 caccagaaggaagaagaatacgccaaataatgsgtcttcttctcaactgagtaacttt 861  
Db 964 GAGAAATGGGGAACAAAGAACAGATTTGTAATGGGTTTCATGATGGCGCCATGTTTGGC 1023  
QY 862 tgcattactcgtctcagcactgccccttttgaaaggttttcgcattgcatcagaaatgac 921  
Db 1024 CTTATGTACTCAACTACGGTCTTGGCTTCTGGATGGGTCTCGTTTCTCGTGTAGATGT 1083  
QY 922 gaggttgcgacgttggcaagctcttactgttgccttcttcgctcaacttagcagccag 981  
Db 1084 CGAGT---CGATGTGGTGATATTCTCACAGTTCTCATGGCCATCTTGATCGGATCGTTC 1140  
QY 982 tccatctcaatgcttgcccttcaggttcaagtcgtttaccacacgcgcacatcttcgctcc 1041  
Db 1141 TCCCTTGGGGAAGCTTAGTCCAAATGCTCAAGCATTTTACAAAGCTGTGGCGCGCGCA 1200  
QY 1042 gaattattcagtatcatgacaaaccccgagctcgacctctcgacctctcgaaag 1101  
Db 1201 AAGATATTGTGNACGATCGATCGCCAGTCCCATTTAGATTCATATTCGAACGAAGGGAAG 1260  
QY 1102 cagcagagggctgctcaggttcaaatgagatccaaacctggcatttgcctaccctcc 1161  
Db 1261 ACGCTCGACCATTTTGGGGCCACATTTGAGTTACGCAATGTCAAGCATATTTACCATCT 1320  
QY 1162 cgaccatctgccaaagtaacttcagagatttcaacttcgaatccagcttggaagacgacg 1221  
Db 1321 AGACCCGAGGTCACCGTCATGAGAGATTTCTCTGTCAATGCCGCTGGAAAAAACAAC 1380  
QY 1222 gccctcgtcgtgcatcaggtagcggcaaaagcacaatggtcgcttacttgaacggtgg 1281  
Db 1381 GCTTTAGTGGCCCTCTGGGCTCTGGAAAAGTAGCGTGGTGGCTGTGGTTGAGCGATTC 1440  
QY 1282 tatctgccagttcggggaggaataattacttgatgggttggaaactgggacaatacgaatgg 1341  
Db 1441 TACATGCCCTGTTCGGGTACGGTTTGTGATGGCCATGACATCAAGGACCTCAATCTC 1500  
QY 1342 aaatggctgagaagccgcatctgcgctcttcaacaggaacactgtgttcttctgtggcaca 1401

Db 1501 CGTGGCTTCGCCAACAGATCTCTTTGTTAGCCAGGAGCCTGTTCTTTTGGCAGCAG 1560  
QY 1402 atcttcagaacatttgcacacgggtttcatgg---atgagaacagagatcgcctcgcgcaa 1458  
Db 1561 ATTTATAGAATAATTAGGCACGGTCTCATCGCACAAATACGAGAAATGAATCCGAGGAT 1620  
QY 1459 aacaaatggagcttgtgcaaaaagcttgcaagccgaacatggcgagctgttcttactaat 1518  
Db 1621 AAGTCCGGGNACTCATCGAAGACGGGCAAAATGGCGAATGCTCATGACTTTATTACT 1680  
QY 1519 gagottccgaacgggttatgagactgaagtgtgcgagagcgcgagccattgagtgaggt 1578  
Db 1681 GCCTTGCTGAAGGTTATGACACCAATCTTGGCAGCGCTGCTTCTCTTTTCAGGTGGC 1740  
QY 1579 caacgacaaagaattgcaatcgacgaagtatcataatcgatcccaagaatcctgttacttc 1638  
Db 1741 CAGAAACAGCGCATTTGCAATCGCCGTCCTTGTAGTGACCCCAAAATCCTGCTCCTG 1800  
QY 1639 gatgaactccagcgccttgcacgaagcgaggaagtggtccagagagcccttgaac 1698  
Db 1801 GATGAAGCTACTTCGGGCTTGGACACAAAATCCGAGCGTGGTTCAAGCAGCTTTGGAG 1860  
QY 1699 cgagttccaaaagccgactacttttggtaattgcccacaaaactagccactgtcaaaagt 1758  
Db 1861 AGGCGAGCTGAAGCGCGAAGTACTTATGTGATCGCTCATCGCTTTCACGATCAAAACG 1920  
QY 1759 gctggcaacatcgacgactcatttccagggggaagaaatcgtagaagacacacacacaa 1818  
Db 1921 GCGCACAAATGTGGTCTGGTCAATGGCAAAATGCTGAACAAGGAACCTCAGCATGAA 1980  
QY 1819 ttgatgaattcggctgctcattacgcgcgactggtggtgac----- 1861  
Db 1981 TTGTTGACCGCGAGCGCTTATCGCAAACTTCTGGAGGCTCAACGTATCAATGAACAG 2040  
QY 1862 -----aggaacctcgggctgacgaacaaacagaaacatgagaagacc 1902  
Db 2041 AAGGAAGCTCACGCTTGGAGGCGCGCGCTCAGGATCTCAGGAATGCAGATATTGCC 2100  
QY 1903 ctgcagaaagcagcagcagagagctgctggtgaacagccgacattgagcgcactcac 1962  
Db 2101 AAAATCAAAATCGCTCAAGCGCATCATCCGATCTCGACGAAAACCCACACCACTTAC 2160  
QY 1963 accactgccacatctcaagctggagacctggagaagc-----ggaaggtgcggtcg 2016  
Db 2161 CGCACGGGACCCACAGTCTGTTCCAGCGCGATTTCTTTTAAAGAGACCCCGGAAACA 2220  
QY 2017 actttgggctact-----cgctcctaaaaatgcataatccttaatggttcttagaaca 2067  
Db 2221 ACTCCGAAATACTCATTTATGGACGCTCTCAAAATTTGTTGCTTCTCAACCGCCTGAA 2280  
QY 2068 aaaaatctctactggtgctctctgttgcacaataaacggtcttgatgtagcgcgccaca 2127  
Db 2281 ATCCGTACATGCTCATCGGTCTGTGTTCTCTCAGTGTAGCTGGTGGTGGCCAAACCCACG 2340  
QY 2128 ttccaggacaagccctttgttttcgagatgctcactgcttctcaggttgaagtggtcat 2187  
Db 2341 CAAGCAGTGTATATGCTAAAGCCATCAGCACACTCTCGCTCCCGAATCACAATATAGC 2400  
QY 2188 gcggaacaggaagcagacactttataatctctgatttcttcttctggtggtcctcaggaat 2247  
Db 2401 AAGCTTCGACATCATCGGATTTCTGCTCATTTGATGTTCTTCGTTGGTGTATCATTCAG 2460  
QY 2248 ctagttagtatattcagattggctgagacatgcaacgttatttccaaagtgtcaccatc 2307  
Db 2461 TTTTATCAGCAGTCAACCAATGGTGTGCAATTTGCCGTATGCTCCGAGAGACTTATTTCGT 2520  
QY 2308 cgatacgaagccgaatgttcccaacagtagtactggatcgaagacatcgaaactcctcagatc 2367  
Db 2521 CGGCGGAGAAGCACTGCCTTTCCGACGATACTCCGTCAAGACATTTGCTTTTGGACAAG 2580  
QY 2368 cgggaacaaatttctgctctcacaatcgcaactgtcagctctacccagcagattgcaa 2427



KSVSAAVSKKVPKPEOFKYSLSLTKVFIKAFNRPGLYMLIGLTFSLAGGQPTQAF  
LYAKAISTLSIPESFHKLRHDANFWSLFFVVGIAQFISLSINGTAFALCSRLIRR  
ARQAFRSILRODLSFFDEENSTGALTSFLSTETKNSLGVSGVLTGTMSTTLGA  
AMIALAGWKALALVCISVYPILLACGFLQQRSSAYEGGSASYACEATSA  
IKTVASLRQDWGVIYHDLQKQGRKSLISVLRSLYSQAALFFVFCVAGLFWYGG  
TLLGHREYSIFRFFVCFSEILFLGASAGTVFSPADMKAKNAQAQFKFLFDSKPTID  
IWSDEKLESMEGIEFDFVHFYRTPRQPSVLRGLNSVVKPGOYIALVSGGCGKS  
TIIALLERFYDALAGGVFDGKDIITKLNVNSRSLNSVSBPTLYOCTIKENILLYG  
DKDDYSEETLIKVCKDANLYDFVMSLPEGFDFVSGKGMSSGQKORVAIRALLRD  
PKVILLDEATSLDSEKVVQALDAARGRITTAVAHRLSTIONADIIYVPDQGI  
VESGTHHEIRNKRYELVNLQSLGKTH\*  
BASE COUNT 967 a 995 c 1031 g 1057 t  
ORIGIN

Query Match 11.9%; Score 450.6; DB 8; Length 4050;  
Best Local Similarity 47.5%; Pred. No. 3.7e-115;  
Matches 1769; Conservative 0; Mismatches 1869; Indels 87; Gaps 11;  
QY 116 gatatgccacaaatcacgactggactctcaatgtcatcgctcactcgccatcgatcgat 175  
DB 302 GCTATGCTCGAAGAGGACATCTCATCTCTGAGCGCAATCTGTGCTATTGCTG 361  
QY 176 ccgggggtccctctctgatgtogatatctcttcttgatcttcacacaaagtccaaca 235  
DB 362 CTGGTGGCGCTCTTCCCTTGTTCACAAATCTCTTCTGCTTCACTAGGCTCAGCGTTCAGG 421  
QY 236 attacaattccggcgacggagtcctgaagcttcaagccgatggatcattctgctc 295  
DB 422 GCATATCTCTGGGTACTATGCCCTATCATGAGTTCATCACAACACTGACTAAGAAATGTGC 481  
QY 296 tdtgggtctacacctctttattgggaagttgtctcactcactgacttccacggctgcca 355  
DB 482 TTTACTTTCTGTATCTCGGTATTGCGGAGTTTGTCTACAGTCTATGTCAGCAGCGTGGT 541  
QY 356 ttacatttcagctatacagaacacactgaactctctgaagagtgtctctgaatgcact 415  
DB 542 TCATTTTACTGGCAACATCTCACACAGAGATCCGTGAAATATATCTTGAGGCTATCC 601  
QY 416 tgcggcaagagtgctggcaatttcgaacagcagagaatggcaatcgccactcargtca 475  
DB 602 TGAGGCAGAAATATGCTTACTTCGACAGTTGGGCGCGGTGAAGTTACACGCGTATCA 661  
QY 476 ctaccaatggcaacgctatacacaacaggattgtccgagaaattgctttacccgtgcagg 535  
DB 662 CTGCTGATACCAACTGATCCAGGAGCGCATCTCTGAGAAAGTTGCTCTCACTTTGACCG 721  
QY 536 cacttcaatgtcttctctgattgtgtgctgttggcgtctcagtggaagctagctt 595  
DB 722 CATTCGCCACATTTGTAACCGCATTTATGTGCGCTACGTCAAGTATTGGAAGTTGGCTC 781  
QY 596 taataccatgtccgtcatccctgcctcatttctctggtcaccggcatctgcatagcaattg 655  
DB 782 TGATCTGTACCTCAACTATCGTCGCGCTGCTCATGTTATGGAGGTGGGTGAGGTTTA 841  
QY 656 atgcgctcagaggccaggaatcaccagatatactatcagccgctgtcctcgcagaag 715  
DB 842 TTGTGAATACAGCAAGAAATCTATTGAAGTTATTTGCTGCTGGTGGAACTGTGCGGGAAG 901  
QY 716 aagttcttatcatcccgacagtcctcattctctcagccacagaagaaatggctcgaaa 775  
DB 902 AAGTCATCAGCTCCATTCGGAAATGCTACCGCTTCGGCACTCAGGATAGCTCGCCAAGC 961  
QY 776 aatgatgtgttttttgcagcaagcacaccacgaaggaaggaagaaatcgccaataatg 835  
DB 962 AATACGAAACCATCTGCTGAGGCTGAAATGSGGCGTCAACACACAGGTCATCCTTG 1021  
QY 836 gsgtctgttctcaactgagtacttttgcattacgtctgatcgactgacctgttttga 895  
DB 1022 GTATGATGATGGGTGATGTTTCGGTATCATGTTCTCGAACTATGCTCTCGGTTTCTGGA 1081  
QY 896 aaggttttcgatgtatcagaatggcgaggttgcgacgcttggcaaaagtcttactgttg 955

DB 1082 TGGGATCTCGTCTCGTTCTCGTAAAGAAGT--CAACGTGGGCCCAAGTCTTCACAGTTT 1138  
QY 956 tcccttcgctacacttagcaccacgctccatctcaatgcttgcgcttcaggttcagtcg 1015  
DB 1139 TGATGTCTATCTGATGCTGTTCTGTTTCTAGTTTGGGCAACGTCGCCCAATGGTCAGGCT 1198  
QY 1016 tttaaccaacgcccgcactcttcgctccgaaattattcagtatcatattgacaaaacccacgacg 1075  
DB 1199 TTACGAATGGTGTGTGTCGCGCCGCGAAGATTTTACAGCACGATTTACCCGACATGCCAC 1258  
QY 1076 tgcacctctcgacccctctctggaagcagcgagggctgcctaggtcgaattgagatcc 1135  
DB 1259 TGGACCCCTATTCTGACGAAGGGAAGGTACTCGACCATTTTGAAGAAATATCGAATTTTC 1318  
QY 1136 aaacctgggaatttgcctacccctcccgaccatctgcgccaagtactctcgagattcaact 1195  
DB 1319 GCAATGTCAAAACATCTATCCCTTCAAGACCGGAAGTTACAGTCATGGAAGATGTCTCTT 1378  
QY 1196 tgacaattccagctggcaagacgacggccctcgtcgttgcgcatcaggtacgagcaaaagca 1255  
DB 1379 TATCGATGCCCGCGCGGAAGACTACCGCATTTGGTGGGCCCATCTCGCTCTGGAAGAGTA 1438  
QY 1256 caatggctgcttacttgaaacggtgtgtatctgcccagttcggggaggatattacttggatg 1315  
DB 1439 CTGTTGTCGGCTTGGTGAGCGCTTTTACCTTCCAGTAGGAGCCAGGTATTGCTGGACG 1498  
QY 1316 ggttggaaactgggacaatacaaatgtgaattggctgagaaagccgcatcgcctcgttcaac 1375  
DB 1499 GCATGATATCCAAACCTCAACCTTCGCTGGCTGCGACAGCAGATCTCTCTTGTCAAGC 1558  
QY 1376 agaaactgtgtgttttcgtggcaaatcttccagaaacttgcgaacggtttcattggatg 1435  
DB 1559 AGGAACCTGTTCTTTCAGCACACAGGATCTTTAGAAACATCGAACATGGCTTGATGGCA 1618  
QY 1436 agaacaag--atctgctctgcgaaacaaatggagcttgtgcacaaagcttgcaag 1492  
DB 1619 CCAATTCGAGCATGAGTCGAAGGACAAAGATTAGAGAGCTGCTTGAGAAATGCGGCCAGAA 1678  
QY 1493 ccagcaatggcagctgttctcattcaatgagcttcgaaagcgttatgagactgaagttggcg 1552  
DB 1679 TGCCCAATGCTCATGATTTTATTATGCTCTGCTCTGAAGGTTAGCATACGAATGTGGGTC 1738  
QY 1553 agcagccggagccttgagtgagggtcaacgcgaacaaatgtgcaatcgcacgaagtaca 1612  
DB 1739 AGCGTGGTCTTCTTACTTTTCAGGAGGTGAGAAGCAACGATTATGCCATTCCTCGTCCATTG 1798  
QY 1613 taccgataccagatccttctactgatgaagctacagcccttgaccgcgaagggcg 1672  
DB 1799 TCAGTACCCCAAGATTTCTGTTGATGAAGCTACATCAAGCTTTGGATACCAAGTCCG 1858  
QY 1673 agaaagtgttccaggagccttgaaacgagtggtccaaagaccgcactacttctgctcattg 1732  
DB 1859 AGGCGTCTGTCGAACGCGCTCTTGATAAAGCTGCCGAGGTAGAACTTACCATTGTCTATTG 1918  
QY 1733 cccacaaactagccactgtcacaagtgtcggcaacatcgcagctcatcttcccaggggaaaa 1792  
DB 1919 CTCACGCTTGTCAACAATCAAAACAGCCCAACATTTGTTGCCATGCTGCGCGGAAAGA 1978  
QY 1793 tgcgtcagaagaagcagcaccacgaattgatcgaattcggcgtgcatcgaagccgactcgg 1852  
DB 1797 TTGCGCAACAGGGAACACATGACGAATTTGCTGATCTCAAGGCACTACTATTAACCTTG 2038  
QY 1853 tgcgtgcacagggacctcggggctgaggaacaaacaaacatgagagacccctgcacgaaa 1912  
DB 2039 TGGAGCGCAACGCTATCAAGGAGGAGAGAGAGAGAGAGCTCTGGAAGCCGACCCGAC 2098  
QY 1913 aggcagcagagaagtgtcgttgaacg-----acoggcacttgagcgcactcaccac 1967  
DB 2099 TGGACCCGATGATTTTGGTCAAGAGGGGTTTACTCGCATCAAGACTGCAGTTAGCAGCT 2158  
QY 1968 tgcacatctcaagc-----tgagagactggagagcgggaaggt 2006  
DB 2159 CGAATTTCTCTCGATGCTGTAGATGAGAAAGCGCGCTTTGGAGATGAGCGAACCGGAACG 2218

QY 2007 gcgggtgggaactttgggtactctgctctcaaaa----- 2040  
Db 2219 AAAAATCGGTTTCGAGTGGCGTTCTATCTAAGAAAGTCCCGAGCAGTTCGAGAAGTATT 2278  
QY 2041 -----Lgcactcaatcatgtctacgaacaaataatctctactgtctctctgtgtg 2095  
Db 2279 CACTCTGGACCCCTCGTCAAGTTTCATGGTGCAATTAACCGCCCTGAGCTTGGTTACATGC 2338  
QY 2096 caacaataacggttcttgatagcgcgccgacatttcacagacaaagccctttgttttcg- 2154  
Db 2339 TCATTGGCTTGAATTCCTGCTTCTGCGGTGGTGGTCAACCTTACCCAGGCTTTTCTGT 2398  
QY 2155 -----agattctcaactgtcttcgaattgagtggtcactcgcgcgcacaggaac 2200  
Db 2399 ACGCCAAGGCAATTAAGCACCTCTCGTTACCCGAATCAATGTTCACAAATCAAGGATG 2458  
QY 2201 ggcagacttttatctctgatgtctcttctgttggtctaggaatactctagtaggattt 2260  
Db 2459 ACGGAATTTCTGCTCTGATGTTCTTCTGTTGGTGAATTTGCTCAATTTATCAGCCTGT 2518  
QY 2261 tcacgattggtgacatgcgaacgttatttcacaagttgtcccccctcgtctatacgaocg 2320  
Db 2519 CTATCAATGTACAGCATTTGCTATTGTTTCGGAGAGACTCATTCGCCGAGCTAGAAATC 2578  
QY 2321 caatgttccaaagactactgatacaagaatcgaactcctcagacatcccggagcaaaattt 2380  
Db 2579 AAGCGTTTAGATCGATTTCTTCGTGACAGACATCTCATTTTCGACGGGAGAGAACAGCA 2638  
QY 2381 ctggtctctcaatgcgaactgcagctctaccacgcaggttcaagagttgatatacag 2440  
Db 2639 CCGGTGCTTGAGCTCTCTTATACACAGACAGCAAGATCTGTCGGCGGTAGCGGAG 2698  
QY 2441 caaatltt-tctcatattatcgttgcgtgcaacatcgtctcgcagcagtgctctacac 2498  
Db 2699 TGACTCTCGGCACAATCATGACTAGCAGTACGCTTGGGGCGCGATGATCATTTGCAT 2758  
QY 2499 tagcctatgatgaaactgggctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2558  
Db 2759 TGGCGATCGGATGGAAGTGGCTGTGTTGCAATTTCTGTGCG---TTCCAAATCCTTCTCG 2815  
QY 2559 tggctggtacctcagaaatcgtctagagacgaagctagaagcgcggaactcgcgaact 2618  
Db 2816 CATGCGGCTTCCTCAGATTTCTACATGCTTGTCTCAATTCACCAACGATCGAAGTCTGCCT 2875  
QY 2619 ttgcagaaagtgtgggcttgcaagcgaagagtttacgcgcagtcgcgcagcgtctctatctt 2678  
Db 2876 ACGAAGGGTCTGCGAGCTATGCTTTCGGAAGCCACGTCAGCGATCCGCACTGTAGCATCAC 2935  
QY 2679 tgactctgaaggccatgtttctcacaacagtaactcgcgcagcagtgtagcaagctcttgcata 2738  
Db 2936 TCATCCGGAACAAGATGCTCTGGGGCGCTTACCACGACCAAGCT--ACAAAAACAGGGAGC 2993  
QY 2739 gatcatcaagctttgttttgacgatgtttgttttcaactgttcacagtcgatgga 2798  
Db 2994 GAAGAGTTTGATCTCAGTGTGAGATCTCTCCCTGCTGTATGCCCTGCGGAGGCAATTCGT 3053  
QY 2799 gtttttggctattgcctcgtggttttattgttcagtcgataattggctcaggtgagtagac 2858  
Db 3054 GTTTTTCTGCTGCTTGGGCTTCTGATGTTGGTGGTACATTTCTAGGCCATCATGAGTA 3113  
QY 2859 gacsaactcaatttatatactctcgtggcgcttttggtttgcggtccaaagcagcagc 2918  
Db 3114 CAGCATCTTCGGCTTCTGCTGCTCTTCTGAGATTTCTTTTGGTGGCAATCAGCCGG 3173  
QY 2919 ccagttatggcttactccacagagttttacaaagctcgcgcgcgcgcgcgcgcgcgcgcgc 2978  
Db 3174 AACTGTCTTCTTTTGGCCCGGACATTTAGGGAAGCAAGATGCGCGTGTCTCAATTCAA 3233  
QY 2979 ctggctcgcgacattgaagcgcacatccgcgaaacgcgagaggaacaaagagagcc 3038  
Db 3234 GAAACTCTTCACAGCAAGCAACCATTTGATCATCTGTCGGATGATGGCGGAGAAATTCGA 3293

QY 3039 agtgggtgatgccctgtcgacctcgagagacattgaattcaggtatcgtcaacgltgattc 3098  
Db 3294 GTCATATGAAGCGGAATTCGAATTCGGGACGCTCCACTTTAGTATCCCAACGGCCGGA 3353  
QY 3099 ggcctcagttctccgcggggtttccatgacaatcagcagcagacaaatttggtagcttatgt 3158  
Db 3354 CGACGCTCTTCTTGAGGACTGAATTTGAGCGTGAAGCCTGGACAATACATATGCCCTTGT 3413  
QY 3159 gggcgctctcggctggaagcaagtcacgttgcctcttgcgcgcgcgcgcgcgcgcgcgcgc 3218  
Db 3414 TGGACCCATCGATCGGTAAGAGCACTACGATTTGCTCTGTGAGCGAATTTATGACGC 3473  
QY 3219 gacctcgcgcgaatttcatttcacacagaaatattgcagaaatgcgcgcgcgcgcgcgc 3278  
Db 3474 ACTTCTGAGGGTCTTCGTTGACGGAAGGACATTTACAAACTCAATGTCAACTCAT 3533  
QY 3279 ccgcgcacatatgtcttgggtccaaacagaaacccacactttaccagggtccggttcgcga 3338  
Db 3534 CGCAGTTTCTCTCCCTTGTGAGCAAGAACTACTCTGTATCAGGGTACCATCAAGGA 3593  
QY 3339 gaatgtgcgttggccctcgaagcgcgaattatcaagaaga-gcttgcgaagcgcgccttc 3397  
Db 3594 AAATATCTGCTTGGAGTCGATAAGGATGACGTTTGGAGGAGACTTTTATTAAAGTCTG 3653  
QY 3398 ccgcgaagcgaatgcttgggatttgcattctcttaccagaagcgttggaaacgccttg 3457  
Db 3654 CAAGATGCCAACATCTATGATTTGTTATGTCATCTCCCTGAGGGATTTGACACCGTCTG 3713  
QY 3458 cggctcaacgagggatgcagttctccgcggcggaacagcagcgagtcgcacatcgcaagag 3517  
Db 3714 TGGCAGCAAG-GGAGCATGTTGCTGTGTGGACAAAAACAGCGTTCGCCATTGCTCGTG 3772  
QY 3518 cattgattcgaataccaaagctgttgccttacttgcagcagcagcgcgcgcgcgcgcgcgcgc 3577  
Db 3773 CCTCTCGCTGACCCCAAGGCTTCTTCTGATGATGTCATCTCCCTGAGGGATTTGACACCGTCTG 3832  
QY 3578 aatcgaagcgtctgttcaagctgcctcgcagtcagcagcagcagcagcagcagcagcagcag 3637  
Db 3833 AATCTGAGAAAGTCTGACAAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3892  
QY 3638 cagtgcgccacgaacttccctcatttcgggaattgtgatgtattttgttggtttggcaacg 3697  
Db 3893 CTGTTTGGCCACCGCTGAGCACCATTCAAAACGCTGATATAATTTACGTTTTCGACCAAG 3952  
QY 3698 ggagaatcgcgaagc 3757  
Db 3953 GCAAGATCTCGAAAAGTGGAAAGCCACACGAGTTGATTTCGAAAACAGGGCCGCTATTACG 4012  
QY 3758 agatg 3762  
Db 4013 AGCTG 4017

## RESULT 7

LOCUS AF173826 4569 bp DNA linear PLN 03-FEB-2000  
DEFINITION *Emmericella nidulans* ABC-transporter (abcB) gene, complete cds.  
ACCESSION AF173826  
VERSION AF173826.1 GI:6856158  
KEYWORDS  
SOURCE *Emmericella nidulans*.  
ORGANISM *Emmericella nidulans*.  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Eurotiales; Trichocomaceae; *Emmericella*.  
REFERENCE  
AUTHORS 1 (bases 1 to 4569)  
TITLE Molecular characterization of ABC-transporter encoding genes in  
Nascimben, A.M., Terenzi, M.F., Goldman, M.H.S. and Goldman, G.H.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 4569)  
AUTHORS Nascimben, A.M., Terenzi, M.F., Goldman, M.H.S. and Goldman, G.H.  
TITLE Direct Submission  
JOURNAL Submitted (29-JUL-1999) Ciencias Farmaceuticas, FCFRP-Universidade









```

QY 3721 caagcggaactacaacgactgagagagagatattacgagatg 3762
|||||
Db 4081 CACAGCGAACTGGTCCAGAAAAGCGCGGTACTACGAGCTG 4122
|||||

RESULT 8
AF071411
LOCUS AF071411 6133 bp DNA linear PLN 25-JUL-2000
DEFINITION Aspergillus nidulans multidrug resistance protein MDR (atrD) gene,
complete cds.
ACCESSION AF071411
VERSION AF071411.1 GI:5456700
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
1 (bases 1 to 6133)
Andrade,A.C., Van Nistelrooy,J.G., Peery,R.B., Skatrud,P.L. and De
Waard,M.A.
The role of ABC transporters from Aspergillus nidulans in
protection against cytotoxic agents and in antibiotic production
Mol. Gen. Genet. 263 (6), 966-977 (2000)
20408170
2 (bases 1 to 6133)
Andrade,A.
Direct submission
Submitted (09-JUN-1998) Phytopathology, Wageningen Agricultural
University, P.O.Box 8025, Wageningen 6700 EE, The Netherlands
LOCATION/Qualifiers
1..6133
/organism="Emericella nidulans"
/strain="W6-096"
/db_xref="taxon:162425"
join(<1332..1736,1788..5071,5129..>5486)
/gene="atrD"
/product="multidrug resistance protein MDR"
<1332..>5486
/gene="atrD"
join(1332..1736,1788..5071,5129..5486)
/gene="atrD"
/note="ABC transporter D"
/codon_start=1
/product="multidrug resistance protein MDR"
/protein_id="AAD43626.1"
/db_xref="GI:5456701"
/translation="MSPLETNPLSPETAMRPAETSTTEQASTPHADEKKILSDLS
APSTTATPADKEHRPKSSSNNAVSVNEVDALIAHLPEDEQVLKTOLEEKVNISF
FLGRYATKMDILLMISTICAIAGAALPLETILFGLSLASTFQIMLYQISDFYD
ELKNVLYFVYLGIGFEVTVYSTVGFIYTGHAHQKIREYLYLESILRONIGYFDKL
AGEVTRITADNLQDISSEKVLTLTALFTVAFIAYVYKWLALICSTIVAL
VLTWGGGSQFIKIKYSSKSLDSYGAGTVAEEVISIRNATAFGQDKLAKQVEVHLE
AEKWTGNQIVMGFMIGAMFLGMYSNYGLGPMGSRFLVDGAVDGDILTVLMATLIG
SFLSGNVSNAOAFNVAATAAKIPGTIDROSLPDYPYNEGTLDHFGHLELRVNH
IYPSRPEVTMEDYSLMPACKTTALVGPSSGSKSTVGLVERFVMPRVFLDGDH
IKDLNRLWRLQISLVSGEPVLFGTITKNIHGLIGTKYENESDKVKRELIENAAKM
ANADFTALPEGTENVQGRFLSSGQKQRIARAVVSDPKILLIDEATSLDSTK
SEGVQAALERAAGRTIVIAHRLSTIKTAHNIIVLVNGKIAEGQTHDELVDGAY
KLVQAQRIHQEKDAEDADAEDLTNADIAKTATSSASDLGKPTTIDRTGTHK
SVSAIILSKRPETTPKYSLWTLKFVSNRNPPEIPYMLGLVSVLAGGQPTQAVL
YAKAISTLSIPESQYSLRHADPWSLMFVVGIIQFTQSTNGAFAVCSERLIRRA
RSTAFRILQDIAFPKKEENSTGALTSFLSTETKHLISGVSGLTILMTSTLGA
IILALAGWLKALCISVVPVLLACGRFYMLAOFQSRSLAYGESANFACEATSSI
RTVASLFRERDVIYHAOLDQAGRTSLISVLSRLSYASSQALVFFCVLFWYGGT
LLGHEGYDIFRFFCFSEILIFGAQAGTGFSPADMGKAKNAAEFRFLPFQKPOIDN
WSEGEKLETVGEIEFRNVHFRYPTRPEQVRLGLDLTVKPGQYVALVGPSCGKST
TIALERFYADIAKSLVDGKDIISKLINISVRSPLSVLSQEPITYQGTIKENILLGIV
EDVPREFLIKAKDANIYDFIMSLPEGFNTVSGSKGMLSGGQKQKVIAARALRDP
KILLIDATSLDSESKNVQOALDAAARGTITIAVHRSTIQAQDVIVYFDQKIV
ESGTHSELVQKRGRTYELNLSQISLQKGH"
BASE COUNT 1507 a 1565 c 1446 g 1615 t
ORIGIN

```

```

Query Match 11.2%; Score 424; DB 8; Length 6133;
Best Local Similarity 47.5%; Pred. No. 1.2e-107;
Matches 1759; Conservative 0; Mismatches 1820; Indels 123; Gaps 12;

QY 175 tccggggcttccctccctgatgcgacatctcttgagtagcttcaccacaaggtcaac 234
|||||
Db 1761 TCAGAAAGTCGGGGTAACTTGTCCAGATCCCTCTTCGGTTCGCTAGCGTGCAG 1820
|||||
QY 235 aattacaattcgggcgagggagtcctgaagcgttcaagggcgagatggtgattctcgc 294
|||||
Db 1821 AGGATAAATGTTATATCAAAATCTCGTACGACGAGTTCATGATGAATGACCA 1880
|||||
QY 295 ctggtgttcgtctaacctcttatttgggaagttgttctcctcaagtttccacgggtgc 354
|||||
Db 1881 CTGTACTCTCGTATACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACT 1940
|||||
QY 355 attaccatttcagctatatacgaacccactcgaactctctcagcagagtgctctt 414
|||||
Db 1941 TTCATCTATACGGGAGAACACGCCACGACGAGATCCGGGAGATTTTACCTTGA 2000
|||||
QY 415 ttgcggcaagaggtgtcgtgatttcgacaagcagagaatggagcaatcgccactcarg 474
|||||
Db 2001 CTGGCGCAGAAACATTTGGCTATTTTGTATAACTCGCTCGCGGGGAAAGTGACC 2060
|||||
QY 475 actaccaatggcaacggtatatacaaacaggtatttgcggagaaattggctctttacc 534
|||||
Db 2061 ACACCCGATACAAACCTTTATCCAGGATGGCATTTTCGGAGAAAGGTTCGGTCT 2120
|||||
QY 535 gcacttccaattgtcttcttctgcatcttggctcgtcttggcgtctcagtggaagctag 594
|||||
Db 2121 GCCCTGGCGACATTCGTGCACAGCATTCATTATCGCCTACGTCAAACTACTGGAA 2180
|||||
QY 595 ttaataccatgtccgttcactccctgccttctccttctcctgggtcgcaggtatgca 654
|||||
Db 2181 CTAATTTGCACTCAACAATTTGTGCCCCCTCTTCTCACACTGGCGGGTGTCTCAG 2240
|||||
QY 655 gatcgccgtcaggagggcaggatcacccaggtactcactcacgcgcgcgtctccagaa 714
|||||
Db 2241 ATCATCAAGTACAGCAAAAGTCGCTTGACAGCTACGCTGACGGCGGCACTGTCG 2300
|||||
QY 715 gaagttcttatactccatccggagacagtccttcttctacgcccaagaagaatggctc 774
|||||
Db 2301 GAGGTCACTAGCTCCATCAGAAATGCCACAGCGTTTGGCACCAAGACAAGCTTCCGA 2360
|||||
QY 775 aaatatgatctcttttgcagcaagcacacccaagaaggaagaagaatacgccaataat 834
|||||
Db 2361 CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGAAACAAAGAACACAGATT 2420
|||||
QY 835 gsgtcttgttctcaactgagtagtcttcttgcatctacgctgctatcgcactggccttt 894
|||||
Db 2421 GGTTCATGATTGGCGGCATGTTTGGCCCTTATGTACTCGAACTACGGGTCTTGGCT 2480
|||||
QY 895 aaaggttttcgcatgtatcagaatggcgaggttgcgcagcttggcgaagttcttactgt 954
|||||
Db 2481 ATGGGTCTCGTTCCTCGTGTAGATGGTGCAGT---CGATGGGTGTATATTTCTACA 2537
|||||
QY 955 gtcccttcgcctacccttagcagcacgcctccatctcaatgtcttgcgccttcaggct 1014
|||||
Db 2538 CTCATGGCCATCTTGATCGGATGTTCTCTCTTGGGGAACGTTAGTCAAAATGCTCA 2597
|||||
QY 1015 gtttacaacgcgcgcacatcttcggtctccgaattattcagtalcatctgcaaaccc 1074
|||||
Db 2598 TTTTACAAACGCTGTGGCGGCGCGCAAAAGATATTTGGAACGATCGATGCCAGTCC 2657
|||||
QY 1075 ctgcacccctctgcaccccttcttgaaacgacgcagagggctccttagtgcataat 1134
|||||
Db 2658 TTAGATCCCATATTTGCAACGAAGGGAAGACGCTCGACCATTTTGAAGGCCACAT 2717
|||||
QY 1135 caaaacctggcatttgcctaccctcccgaccatctgcaccaagctactcagatttcaac 1194
|||||
Db 2718 CGCAATGTCAAGCATATTTTACCCTATCTAGACCCGAGGTCAACGTCATGGAGGAT 2777
|||||

```









Db 583 CCTCTACTCTCTATCTTGAAATCGCCCAATTTATTCTCTCTATGTATATCCACTGTGGG 642  
Qy 354 cattacatttcaagtatacgaaccactcgaactcttgcagagagtggtcccttgaatgaac 413  
Db 643 TTTTCATCTACGTGCGGTGAGCAGATCACCCAGAAAGATAGTCGGAATATCTCCATGCCAT 702  
Qy 414 ctgcggcgaagaggtctgtgacatttcgacaagcagagcaaatggagcaaatcgcccaactcagt 473  
Db 703 TCTAGCCAGAACATCGGTTCTTCGATAAATCTGGGCTGGAGAACTTACCACTCGGAT 762  
Qy 474 cactacaaatggcaacogtatatacaaacaggtattgcocagaaaattggtctttaccogtga 533  
Db 763 CACCGCAGACGAGATCTATCCAAAGATGAATATCAGAAAGAGTGGCGCTTGACTCTAAC 822  
Qy 534 ggeactttcaatgtttcttctgatttgcatttgcatttgcatttgcatttgcatttgcatttgc 593  
Db 823 TGCCCTATTCGACCTTCTTCACGCGCTTTATTATCGGATACGTCAGATCTGGAAAGCTTGC 882  
Qy 594 tttaatoaccatgtcgttcacccctgccattttctcgtgtcacccgcatctgcatagcaat 653  
Db 883 ACTGATCTGCAGCTCGACCATGTGTCCGATGATTTAGTTATGGGTGGTATATCTAGGTT 942  
Qy 654 tgatgcogctcagggagccaggtatcaccaggtatctactcaocgcgcgtgtcctcgcaga 713  
Db 943 CGTGTCAAGTCTGGAGGATGACGTTGTGAAGCTATGCTGAAGGAGGACACTGTTGCAGA 1002  
Qy 714 agaagtttatacatccatcccgagacagtccatgcttcttctacgcccagaagaataatggtcga 773  
Db 1003 AGAAGTATCATGCTATATCAAAATPGCCACTGCTCTTGGAAACCCCAAGAGAAATCGCAAG 1062  
Qy 774 aaaaatgatgtcttttgcagcaagcacaccaagaaggaagaagaataatcccaataa 833  
Db 1063 GCAGTACGAAGTGCATTTAAAGGAAGCGAGAAATGCGGGTCGACGGCTTCAAAATGATGTT 1122  
Qy 834 tggsgtcttcttcaactcagtagtacttttgcatttactcgtctatcgcactggccttttg 893  
Db 1123 GGGCATCATGTTGGCTCCATGATGGCTATCATGCTATCATCAAACTACGGGTCTCGGATTTG 1182  
Qy 894 gaaaggttttcgcatgatacgaatggcgaggttgcgcagcgttggcgaagtcttactgt 953  
Db 1183 GATGGGATCACGTTTCTTGTGCGGGGTGAACACACCTGTGTGCCA---TCGTCAACAT 1239  
Qy 954 tgcctttccgtcaccttagcagccacgtcccatctcaatgctgcccctcaggttccagt 1013  
Db 1240 CCTCTTCGGATGTTATTCGGATCTTTAGCATTTGGAAACGTTGTCTCTAATACCCAGC 1299  
Qy 1014 cgtttaccaacgcgcattcttgcgtccgaattattcagtatcataggacaaacccacgca 1073  
Db 1300 TTTTGCATCTGCTATTTCTGCGGTGCTAAGATCTTTAGCACCATCGACCGTGTTCGCG 1359  
Qy 1074 gctcgaccctcgcaccctctgaaaagcagcagcagagggctccttaggtcgaataggat 1133  
Db 1360 AATTGACCCAGGTTCTGATAGGGAGATACCATCAAAACGTTGGAGGGAACCATCGAAT 1419  
Qy 1134 caaaacccctgcatctgacctccctccgaccatctgcaccaagtacttcgagatttcaa 1193  
Db 1420 CCGTGGTATTAGCATATCTACCCATCCGCTCTCTCAAGTTGTTGTCATGGAAAGATATCAA 1479  
Qy 1194 ctggacaattccagctggcagacgacgcccctcgtcgtgcatcaggttagcgagcgaag 1253  
Db 1480 CTTTGTGTTCCAAAGGCAAAACACATGCTCTTGTGGCCCTCAGGTTCCGGAAGAG 1539  
Qy 1254 caaatggtcgttacttgaacggtgttatctgcgcaggttgcgggagagatattacttga 1313  
Db 1540 TACCGTTGTGGGTTGTTGGAGCGATTTCTATAACCCAGTTCTTGGTTCCGTTCTGTGA 1599  
Qy 1314 tgggttgaactgggacatacaatgtgaatggctgagaagccgcatctcgcgttca 1373  
Db 1600 TGGCGGTGATATCAAGACCCCTCAACTCGGATGGCTCCGTCAGCAAAATTTCCCTTTGTGAG 1659  
Qy 1374 acaggaacctgtgttcttctggtggcacaatcttccagaacattgccaacggttctatga 1433  
Db 1660 CCAAGAGCCTACGTTGTTTCGGAACCACTATTTTGTGAACACATCCGATTTAGGTTGG 1719

Qy 1434 tgagcaacgagatctgctcgcgcgaatacaaat---ggagcttgtgcataaagcttca 1490  
Db 1720 GTCCCCAATGGAAATATGAATCCGAGGAGCAATCAAGAACGGATCTCTAGTGGGCCAA 1779  
Qy 1491 agcagcaatggcgactgttcaatagcttccgaacgggttatgagactgaagtgg 1550  
Db 1780 GGAGGCCAATGCTCATGACTTCATTTATGGGCTGCCGATGGGTATGCAACGGACGTCG 1839  
Qy 1551 cgagcgacggcgcccttgagtgagtgtaacgacacaaagaaatgcaatcgacgaagat 1610  
Db 1840 TCAGCGTGGCTTCTTTATTCGGTGCCAGAACGACGATTTGCTATTTCGCCGTGCCAT 1899  
Qy 1611 catatcggtatcccaagatcctgttactcgtatgaagctaccagcgcccttgaccgaagc 1670  
Db 1900 TGTAACGACCCCTTAAGATATTTGCTGTGATGAAGCCACGCTGCTCTCGATACCAAGTC 1959  
Qy 1671 ggagaagtggtccaggagccttgaacagagtgctcaagagaccgactactttggtcat 1730  
Db 1960 TGAGGGGTAGTACAAGCCGCTCTTGATGCTGCTCCGAGGAAGAACCACTATTGTCTAT 2019  
Qy 1731 tgcccaaacactagccactgtcaaaagtgctggaacacatcgcagtcattttcccaaggaa 1790  
Db 2020 CGCCCAACCGTTTATCCACCAATTAATCTGCCGATAATTTGTTGTGATTGTGCGTGGTCG 2079  
Qy 1791 aatcgtcgagcagcacacacacgaattgatcgaaattcggtcgtcattacgcgcact 1850  
Db 2080 CATCGGTGAACAGGTACTCTGATGAGCTGTTGTGATGAAGAGGACATACCTTCAGCT 2139  
Qy 1851 ggtcgtgcacaggaacctgggtgacgaacaacaagaacatgagaagacccctgcacga 1910  
Db 2140 CGTTGAGGCTCAAAAGATCAATGAGGAGCGGCTGAAGAGTCTGAGGATGAGGCTGTGCT 2199  
Qy 1911 aaaggcagcacgagagctgctgtggaacgacggcacttgagcgactcaacacactgc 1970  
Db 2200 CGAAAAAGGAGAGAAATCTCGGCCAGATTTCTGTCTCTGCAAAAGAGTGTCAACTGTG 2259  
Qy 1971 cacatctcaagctggagacctggagaagcggaaaggtccgggtcgggactttgggctac-- 2028  
Db 2260 AAAATACCTGTGACGAGATGTTGAAGCCAAATCTTGGAGGAATTCATACCAAGAGTCCCT 2319  
Qy 2029 -tcgtctcaaatgcatactcaatcatgttctacgaacaaaaaatctactcgtggtc 2087  
Db 2320 TTCGAGCGTTATTTGTCGCCAAAAACGTAGCCAAAGAGAAAGAACCGAATATCGCTGG 2379  
Qy 2088 cttgttgtcaacaataacgg-----ttctgatatgcgc 2120  
Db 2380 CACTTTGATACGATTCATCGCTGGATTTCAACAGCCCGAGAGATTTGATAATGCTCTGTG 2439  
Qy 2121 ggccacatttccaggaacacccctttgttttcgagatgctcactgcttcgagttgag 2180  
Db 2440 ATTCTCTTTTGGCGCTTCTTTGTTGGTGGCGCACCCCGTTCAATCCGCTCTTTTGGCAA 2499  
Qy 2181 tggctcagtcggca-----caggaaacggcgaga 2207  
Db 2500 GGGTATCACAAACATTTTCGCTCCGCCCATCATTTGACGGTAAACTGCGCGAAGATGCCAA 2559  
Qy 2208 ctttttatctctgagtcttcttctggtctctaggaatctctagtagatatttcaagat 2267  
Db 2560 TTTCTGCTCCCTTATGTTTAAATGTTGGGTCTTGTGCCAACTGTTTACCCAAATCCGCGCA 2619  
Qy 2268 tggctggacaatgcaaggtattttcacaagttgtcacccatcgctatacaagcgcgaatt 2327  
Db 2620 AGGTGTTATCTTTGCTATATGTTCTGAGAGTTTGATTTACGGACACGACGAAGTCCCT 2679  
Qy 2328 ccaacgagtagctggatcaagacatcgaaactcctcgacatcccgagacaaaatttctggtgc 2387  
Db 2680 CCGTGCATGCTTCGACAGATATTTGCTTTCGACCTCCCTGACACACACTGGTGC 2739  
Qy 2388 tctcaatcgcaactgctcagctctaccacgcaggttgcaagagttgtatatacgcaattt 2447  
Db 2740 ATTAACGTCATTTCTCTCAACTGAAACCAAGCATCTTTTCTGCGCTGAGTGGTGCCACTCT 2799









```

Db 2667 CTTGCCGCTGGTGGTCAACCTACCCAGGCTTTTCTGTAGCCAAAGCAATATTACGCGCTC 2726
Qy 2164 actgtcttcagattgagtggtcgatcggaacggaacggcgagacttttatctatctgtg 2223
Db 2727 TCGTTACCGCAATCAATGTTTCACAAACTCAGGCATGACGCCGAATTCCTGGCTTGATG 2786
Qy 2224 ttcttctgctggtctaggaataatagtaggatatattcaacgattggtggacatgaac 2283
Db 2787 TTTCTTGGTGGTGGAAATTTGCTCAATTTATCAGCCCTGTCTATCAATGGTACAGCAATTTGCT 2846
Qy 2284 gttatttcacaagttgtcacccatcgctatcaagcgcgaatgttccaacgagtaactggat 2343
Db 2847 ATTTGTTCCGAGAGACTCATTCGCCGAGCTAGAACTCAAGCTTTAGATCGATTCTTCCT 2906
Qy 2344 caagacatgaactcctcagacatcccgagagcaaatcttctggtctctcacaatcgcaactg 2403
Db 2907 CAGGACATCTCATTTTTCGACAGGGAAGAAAGAACAGCAGCGTGCCCTTGACGCTCTCTTTA 2966
Qy 2404 tcaagctctaccacacagttgcaagagttgatcagcaaaattttctcatttatcgctt 2463
Db 2967 TCAACAGAGACGAAGAATCTGTGCGGGCTTACGGGAGTGACTCTCGGCACAAATCATCATG 3026
Qy 2464 gtccgg--tcaacatcgtctcagcagtgctctaccactagcctatggatggaaactgggc 2521
Db 3027 ACTAGCACTACGCTTGGGCGCGCATGATCATTTGCATTGGCGATCGGATGGAAGCTGGCT 3086
Qy 2522 ctggtggttggttgggtgcaacttccaccctgcttggcttggctacactcagaattcgt 2581
Db 3087 CTGGTTCGATTTCTGTGCG---TTCCAATCTTCTTGGCATGCGGCTTCTCCAGATTCTTAC 3143
Qy 2582 ctagagagaagctagaagcggaaactcgcgaactttgcagaagtgctggcttgca 2641
Db 3144 ATGCTTGCTCAATTCACGACGATCGAAGTCTGGCTACGAGGGTCTCGGAGCTATGCT 3203
Qy 2642 agcgaagcagttaccgcgacatcgggaccgtctcatcttttgactctcgaagggccattctc 2701
Db 3204 TCGAAGACCACTGACGATCGCATGCGACTGTAGCATCACTCACTCGCGCAACAAGATGCTGG 3263
Qy 2702 caacagtaactcggacatgttgagaaggtcttgtaagataccaaagctttggttg 2761
Db 3264 GCGGTTTACCAGACACAGCT--ACAAAACAGGGACGGAAGATTGTGATCTCAGTGCTGA 3321
Qy 2762 gacgatgtttgtttcacttgtcacagtgcagtgaggttttggctattggccctgggat 2821
Db 3322 GATCTCTCCCTGCTGTATGCTTCGTCGAGGCATTTGGTGTTCCTTCGCTGCTTTGGGCT 3381
Qy 2822 ttgttatgcaagtcgataaattggtcttcaggtgagtcacacaaactcaatttatcat 2881
Db 3382 TCTGATATGGTGTACACTTCTAGGCCATCATAGAGTACAGCATCTTCCGCTTCTTCGTCT 3441
Qy 2882 ctctgtggcgtttgtttgttcggttccaaagcagccagcagatttggcttactccaag 2941
Db 3442 GCTTTTCTGAGATTTCTTTTGGTCGCAATACAGCGGAACTGCTCTTCTCTTTGCCCCGG 3501
Qy 2942 gttttaccaaagctcgtcgcgtcgaaactatactctctggtcgcggacattgaagccga 3001
Db 3502 ACATGGGTAAAGCAAGAATCGCGTGTCTCAATTCAAGAAACTCTCGACAGCAAGCCAA 3561
Qy 3002 coactccgcgaacgagagagaacaaaaaggccagtggttggtatccctctgtcgacc 3061
Db 3562 CCATTGACATCTGTGCGGATAGGGCGAGAAGTTGGAGTCTATGGAAGCGGAATCGAAT 3621
Qy 3062 tcgagagacattgaattcaggtgtcgaactgattcggctgattcggctgagttcccggggttt 3121
Db 3622 TCCGGGAGCTCCACTTTTAGTACCACAAAGCGCGGAGCAGGCTGTCTTCGAGGACTGA 3681
Qy 3122 ccatgacaatcagcaggacaaatttgatgttattggtggcgtctcttggtctggtggcaagt 3181
Db 3682 ATTTGAGCGTGAAGCTGGACAATACATTGCCCTTGTGGACCCAGTGGATCGGTAAGA 3741
Qy 3182 caacgttgatcgtctgtcggaacgatttctacgaccccgacactcgggcccgaatttcat 3241
```

```

Db 3742 GCACTAGGATTGCTGCTGCTTGAAGCAATTTATATGACGCACTTGTCTGAGGGGTCTTCGTTG 3801
Qy 3242 cacacagagaataattgcagaaatgtcgccgcgttggtaccgcggccatattgctttgggtcc 3301
Db 3802 ACGAAAGGACATTTACCAAACTCAATGTCAACTCATACCGCAGTTTCTCTCCCTTGTCA 3861
Qy 3302 aacaggaacacacactttaccagagctccgttcgcgagaaatgtgacgtttggccctcgaag 3361
Db 3862 GCCAAGAACCTACTCTCTATCAGGGTACCATCAAGGAAATATCTCTGTTGAGTCCGATA 3921
Qy 3362 cogaattatcagaagagctttgttc-aagagcgccctcccgcaagggccaaatctttggatt 3420
Db 3922 AGGATACGCTTTGGGAGGAGACTTTGATTAAAGTCTCAAAGATGCCAACAATCTATGATT 3981
Qy 3421 ttgtaactctttaccagaag----- 3441
Db 3982 TCCTTATGTACCTCCCGTAAGTTTACTGCTCGTGACTCTCTTTCATGCTCGTGGGGTCTT 4041
Qy 3442 -----gctttgaaacgccttggctcgaacgagggatgca-gtt 3479
Db 4042 GCTAACATGCTACAACAGTGAAGGATTTGACACCGTCTGTTGGCAGCAAGGAGGCATGTT 4101
Qy 3480 ctccggcggaacacagcagcggtatcgccatcgcaagagcattgattcgaataatccaaagct 3539
Db 4102 GTCTGTTGGACAAAACAGCGTGTGCGCATTTGCTGTGCCCTCGCTGACCCCAAGGT 4161
Qy 3540 gtgtcacttgcaagcagcagcgtcagccctcgacacgaactcgggaagctctggttccaagc 3599
Db 4162 CTTCTCTCGATGAACCCACATCTGCTTCTGACTCCGAATCTGAGAAAGTCGTACAAGC 4221
Qy 3600 tgccttcgatgagcgatccacagcagcgcgaacgaataagcagtggtgcccagcacttccac 3659
Db 4222 TGGCTTGGATGCTGTGCCCGCGGCGACACAGATTGCTGTGCCACCGCTGAGCAC 4281
Qy 3660 tattcgaatgttgatgttatttttggtttgcacagggagaaatcgccgaagggggcac 3719
Db 4282 CATTTAAAACGCTGATATAATTTACGTTGTCGACCAAGGCAAGATCTGCGAAAGTGGAA 4341
Qy 3720 tcacgggaactacaacgactgagaggaagatatattacgagatg 3762
Db 4342 GCACCAGGATTGATTCGAACAAGGCCGCTATTACGAGCTG 4384
```

## RESULT 12

AR013812 LOCUS AR013812 3924 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 1 from patent US 5773214.

ACCESSION AR013812  
VERSION AR013812.1 GI:3971266

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

UNCLASSIFIED.

REFERENCE 1 (bases 1 to 3924)

AUTHORS Peery, R.B. and Skatrud, P.L.

TITLE Multiple drug resistance gene of *aspergillus flavus*

JOURNAL Patent: US 5773214-A 1 30-JUN-1998;

FEATURES Location/Qualifiers

source

1..3924

/organism="unknown"

BASE COUNT 915 a 1018 c 1021 g 970 t

ORIGIN

Query Match 9.5%; Score 362; DB 6; Length 3924;

Best Local Similarity 46.8%; Pred. No. 3.3e-90;

Matches 1752; Conservative 2; Mismatches 1892; Indels 116; Gaps 14;

Qy 97 gactatctacgaattcttcagatatgcccgaacaaatcacgactggacttcaatgtcatcg 156

Db 136 GGCTTTTTCGTATCTATCGATATCGACATCGATGGGATGTCGCATTTCTCTCGGCAGT 195

Qy 157 ctcatctgcccactcgatcccggggttcctctctctgtatgtcgatcatcttggtagc 216

Db 196 GCGCTTGTGCGATCGCAGCGGTCGAGCACATACCTCTTTCACGGTCTCTTCGCGCAGG 255  
QY 217 ttcaacaagaagttaacaattacaattcgggcgcagggagtcctgaagcgttcaaggcc 276  
Db 256 TTGAGCTCGACCTTTCAAGACATCGCTACCCATCCGATTAAGTACGATCACTTCCATCAC 315  
QY 277 gatggatcatttgcctgctggttgcctgctacacctttatttgggaagtttgcctcaag 336  
Db 316 GAGCTGACCAAGAAATGTGGTCTACTATATATCTTGGGTGCTGCTGAGTTCGTGCGGAT 375  
QY 337 tacggtttcacagcgttcattaccatttcagctatcacgaacacctgaaactcttcgacga 396  
Db 376 TATCTAGCAGCGTCGGCTTCATTTATACCGCGGACCATGCTGACAGAGATTCGGGTG 435  
QY 397 gtgttccttgaatgcacettgvggcaagggctctggaatttgcagaaagcaagcaatga 456  
Db 436 GAATACTTCCAGGCTATCTTAAGGAGAAATATGGCTTCTTCGACACTCTTGGTGTGGG 495  
QY 457 gcaatgcacactcargtcaactaccaatggcaacgctatacaaacaggatttcgcaaaaa 516  
Db 496 GAGATCACAAACGAGATCACCCGACAGCAACCTGATTCAGACGGTATCTCCGAAAG 555  
QY 517 ttggtctttacagtcgaggaacttcaatgtttcttttctgcttgcatttggctgcgtttggcg 576  
Db 556 GTCGGCTTAGCCTTAGCCGGCTTCGAGCTTCGTGACGCGTTCATCATCGCTATATA 615  
QY 577 tctcagtggaagctagctttaaoccatgtctcgtcgcacgcttgcacgcttgcacgct 636  
Db 616 AAGAACTGGAAGCTGGCAGACATCTGTAGTGCAGGCTGCTGCGATTTGCTTGACCATG 675  
QY 637 gcatctcagatagaatgtatgcgctcagaggcagcagatcacacaggatctactcaagc 696  
Db 676 GTGGTGTCTCCACCTCATGCTCAATTTTACGAAGAAAGCATTTGGAATACAGGTCGT 735  
QY 697 ggcgctgtctcgcagaagaagcttatacctatccatccgcagacgcatgcttcttcacgc 756  
Db 736 GCGCCAGTATGCGCGAGGACATCTTGACTCCATCCGACCGCTGCGCATTTGCTTAATGCT 795  
QY 757 cagaagaagaatgctcgaaaaaatgatgtcttcttttgcagcaagcacacaaagaaggaag 816  
Db 796 CAGGAACACTGGCGGGAAGTACGAATCGCATCTCAAGGATGCCGAAGGACCCGGAATG 855  
QY 817 aagaatgcgaataatgsgctgttctcgaactgagtaacttttgcatttaacgctgct 876  
Db 856 AAGTCNAAGTGATCTTCGCCATCATGGTGGGAGCAATTAATGCAATCATGTATCTCAAC 915  
QY 877 atcgactggccttttggaaagtttgcgatgtaacagaatggcgaggtt---gcccagc 933  
Db 916 TATGGACTTGGCTTCTGGATGGGTCTCGCTTCCCTCGTCAAGGAATCAGTAACATCAA 975  
QY 934 gttgcaaaagtcttactgtgttccttcctgcacacttcagccagcagctccatcaatg 993  
Db 976 GCGGGGATGTGCTGACCATCATGATGGCATCATCTAGTTTCTTATATCTGGGCAAT 1035  
QY 994 ctgagccttcaggttcagctgtttaccacgcgcgcatctcggtcccaattatcagt 1053  
Db 1036 GTGGCCCCAATGGACAAGCGTGTGCGAGCGTGTGGGCGGCGACATCAAGCTGTATGGC 1095  
QY 1054 atcattgacaaacccacgagctcgaccctctcgaccttcttgaccttctggaagcagcagagggc 1113  
Db 1096 ACCATTGACCGTCAATCCCTTTAGATGGTGTGTCCGATCAGGCGAAGAGCGTGGAGTTT 1155  
QY 1114 tgcctaggttcaaatggatgatacaaaacctggcaatttgcataccctcccgaccatctgcc 1173  
Db 1156 GTCCGAGAAATATTACTTCCAGAAACATCAGGACAGCTTTATCCCTCGCAGCCGAGGTC 1215  
QY 1174 caagtactcgagatttcaacttgacaattcagctggcaagacgagcgccctcgtcggt 1233  
Db 1216 ACTGTGGCCATGATCTGATGTGCTATATTCCTGCGCGGTAAAGACGACTGCTCTCGTGGGT 1275  
QY 1234 gcatcaggtagcggcaaaagcacaatggtcggtcttacttgaacggtgglatctgcccagt 1293

Db 1276 CCCTCAGGGTCCGCAAAAGTACCATCATCAGCTTGCTTGAGCGGTTTTTATGACCCAGTC 1335  
QY 1294 tccgggaggaattactgtatggttggaaactggggaacaatacaatgtgaatgactgaaga 1353  
Db 1336 GCGGGGACCAATTAATGCTCGACGCGCCACGATATCCAGACATTTGAATCTCCGATGGCTGGA 1395  
QY 1354 agcgcattcgctcgttcaacaggaacctgtgtgttctggtggcaaatcttccagaac 1413  
Db 1396 CAGCAATGTGCTCGTTCGACAGAGCCGAGACTCTTCGCCACCACTCCGGAGAA 1455  
QY 1414 attgcaaacggttttcattgtagcaacagatctgctcgcgaaaaaaataa---ggag 1470  
Db 1456 ATCCGTTATGTATCATATTGGCTCTCGATTTCGAAAAGGAATCGAGTACGAGATCGAAAG 1515  
QY 1471 ctgtgcaaaaagcttcaaaagccagcaatggcgacgtgttccatttaagacttccgaac 1530  
Db 1516 CGAGTCGAAGCAGCAGCTGAATGGCAATGCGCACGACTTCATCATGGCGCTTCCGAA 1575  
QY 1531 ggttatgagactgaagtggcgagcgagcgagccttggagggttcaacgacaacga 1590  
Db 1576 GGTATGATACAAATATCG-----AGAGCTTCTCGCTTCTGGGGCCAGAGCAGCGC 1629  
QY 1591 attgcaatcgcaagaagtaataatcgatccccaaagatcctgttactcgatgaagctacc 1650  
Db 1630 ATCGCTATTGCCAGAGCTATCATAAAGACCTTAAGATCCCTACTCTTAGAGGAAGCAACA 1689  
QY 1651 aggccttgcacgaagggagaaagtggccagagagccttgaacgagtggtccaaa 1710  
Db 1690 TCGGCTTGGACAGCAATCCGAAAGCTGTTTTCAGGAGCGCTTGCAAGGCTTCTTAA 1749  
QY 1711 gacgcactcttggctcattggcccaaaactagcactgtcaaaagtgtcggcaacatc 1770  
Db 1750 GGTAGACGACTATTGTATCGCTCATCGGCTGTCCACAATCCAGAGGCATACAAATC 1809  
QY 1771 gcaagtcattccccaggggaaaaatcgtcgaagcaagcacacacacgaattgatcgaatc 1830  
Db 1810 ATTGTTCTTGCCAAATGGTCAAGATTGTTGAGCAAGGCGCGCATGAGCATTTGATGGATCGA 1869  
QY 1831 ggtgtcattacgcgcactggtggtgcacagg----- 1864  
Db 1870 CGGGGATCTATTGTGATCATGGTGAAGCGCACGAGATTAGAAACGGTACTCGAGATAC 1929  
QY 1865 -----acctcgggctgcagaaacaagaacatgagaagacccctgcac 1908  
Db 1930 TCCAAGAGATATTCACGCTATTGACGAACTTGTCCCTTAAGCATACCAATGACCTTC 1989  
QY 1909 gaaaaggcagcagagagctgctggtgaaacgcgcgcgacttgagcgcaactcacaccact 1968  
Db 1990 TTCTTCGACAAAGATTACCCGGCGATGACGAGTCCGATATTTACAGCATCTCTCAGTGAT 2049  
QY 1969 gcaacatctcaagctgga-----gac 1989  
Db 2050 GACGCATCCGATATTGGATTGCACACTGGGAGAAACAAGACCAGTGTGAGAAATGTCA 2109  
QY 1990 ctggagaagcggaaggtgcggttcgggactttggctactcgtcgtcctctctctctctct 2049  
Db 2110 CTATCTCATCTCAGCAGCGGTCAAGGAAGGCATATTCTTCTGGACTTTGTTCAAG 2169  
QY 2050 atcatgttctacgaacaaaaaaatctctactggtgcttctgtgtcacaataaacggtt 2109  
Db 2170 TTCCTTGGCTCGTTCAACCGTCCGAGTGGCGGTTTTTATTACTAGGGCTCTGCGCATCT 2229  
QY 2110 ctgatatgcgcggccacatttccaggacaagccctttgttttcagattgctcactgctc 2169  
Db 2230 ATACTGGCAGGGGGATCCCAACCATFCGAGGCTGTGCTGTGTTCGAAGGCTGTCAAGCACC 2289  
QY 2170 ttcaggttgaagtgcctg-----cgacacaggaacgagcagacttttatctctg 2220  
Db 2290 CTCAGCTGCCCCCTCTTGAATATCCGAAGCTCCGGCAGATGCCAACTTCTGGTGCCTC 2349  
QY 2221 atgttcttctgctggtcttaggaaatctagtagataatttcaogatttggctggaatgc 2280  
Db 2350 ATGTTTCTCATGATTGGCATTGCTCGCTTGTGTTGTATAGTGTGCAAGGGACCTTGTGTT 2409



ONIRHVPSPEVTVVADLSCYIPAGKTTAFVPSGSKSTLISLLERFVDPVAGTIM  
LDGHDIGTLLNLRQOMSLVSQEPRLFPATFIAGNIRYGIIGSRFERESTVEIRKVE  
AAARMANAHOPIMALPNYGDNIIESFSLSGQKQRIARAIIDPKLILLDEATFAL  
DTSEKIVAAALDASGRRTIVIAHRLSTQKAYNIIVLANGOIVEGQPHHLMDRR  
GIYCDMVEAHEIKRYSKRYSHQNTLSPKRNPMPTFFDKDYPGDESDIYSILS  
DDASDGLHTEKORPVSRLMSLQMPVKEAYSFWTFLKFLASFNRPEWPFLLGL  
CASTLAGIOPSQAVLFAKAVSTLSLPLEYPKLRHDANFWMFLMIGIVSLVLYS  
QGTFLFASISKMVYRARSQAPRVILHODISFFDQOENTTGALTATLSAGTRELGISG  
VTLGTILIVSNLVASIGVALVIGKIALVCISAVPALLMCGFVRVWMLERFORAKK  
AYOSASACEAASAIKRTVSLTNETALQSYQALRQLKSDILPIVKSLLYASSQ  
ALPFCMALGFEGSGLHGEYSLFQFYVCFSEVIFGAQAAGTFFSHPMDMGAKHA  
AREPKRLSSDTMHASRKGVPVTSRMLGVDFRVPYPSRLEOPIRHLNLITKPG  
OFVALGASGSKSTTIALLERFYDPLKGGVYVDGKNITITLEMSSYRSHALISQET  
LFQGTIRENILLGSNTPHVTDFFLVKACKDANIYDFILSLPQGTWIVNGGMLSGG  
OKQRIATARILRPKLLLDDEATLSALOSESEKVVQAAADRAARGRTTIAVAHRLSTI  
ORADLIYVLOGEVESGTREHLLRKKGRYVELVHLONPDATGTR\*

BASE COUNT 915 a 1018 c 1021 g 970 t  
ORIGIN

Query Match 9.5%; Score 362; DB 8; Length 3924;  
Best Local Similarity 46.6%; Pred. No. 3.3e-90;  
Matches 1752; Conservative 2; Mismatches 1892; Indels 116; Gaps 14;

QY 97 gactatctacgaatctcagatgacgacaaatagactggactctcaatgcatcgcg 156  
DB 136 GGCCTTTTGGGTATCTATCGATATGCGACTCGATGGGATGTCGCCATTTCTCTCGGCAGT 195  
QY 157 ctcatctcgccatcgatcggggctcccttcctctgatctcgatcatctcttcgtagc 216  
DB 196 GCGCTTGCTGCGATCGCAGCGGTGCAGCACATACCTCTCTTACGGTCTCTTCGGCAGG 255  
QY 217 ttacacaaagcttcaacaattacaattcggcgacggagtcctgaagcgttcaaggcc 276  
DB 256 TTGACGTGCACCTTTCAAGACATCGCTACCCATCGGATTTACTTACGATCACCTTCCATCAC 315  
QY 277 gatggatcattctgctcgtggtctcctctacctcttattgggagttgtccctcaag 336  
DB 316 GAGCTGACCAAGATGGTCTACTTCATATACTTGGGTGCTGCTGAGTTCGTTCGCGATT 375  
QY 337 tacgtttccacggctgcattaccatttcagctatatacgaacaaactcgaactctcgacga 396  
DB 376 TATCTAGCGACCGTGGCTTCATTTATACCGCGACCATGCTGTACAGACAGATTTCGGGTG 435  
QY 397 gttccttgatgaatgcaccttcggcgaaggtgtggtcatttcgaagaagacgaatgga 456  
DB 436 GAATACTTCCAGCTATCTTAAGGCAAGATATGCGCTTCTGTGACACACTCTTTGGTGTGGG 495  
QY 457 gcaatcgccactcargctactacaaatggcaacccgtatacaaacaggtattgcccgaaga 516  
DB 496 GAGATCACAAACGGGATCACCGCGACACGAACTGATTCAGACGGTATCTCCGAAAG 555  
QY 517 ttggcttttacogtgcaggaacttcaatgtttcttcttgatcttggcttggcttggcg 576  
DB 556 GTCGGCTAGCTTGCACCGGGCTTTCGACGTTGCTGACGGCTTTCATCATCGCTATATA 615  
QY 577 tctcadtgaagtagctttaataaccatttcgctatccctcctcatttccctggctacc 636  
DB 616 AAGAACTGAAGCTGGACACTTAATCTGTAGTGCACGCTGCTGACATTCGTTGACCATG 675  
QY 637 ggcattcgatagcaattgatgcgctcagaggccaggatcaccaggatctactcagc 696  
DB 676 GGTGGTTGCTCCACCTCATGCTCATTTTATGCAAGAAACCAATTGGAATACCGGGTCTG 735  
QY 697 gccgtgtctctcagaagaagtcttataccatccatccggacagtcattgttttctacgcc 756  
DB 736 GCGCCAGTATGGCCGAGGACATCTTGACTCCATCCGACCGCTCGCAGCGTTTAATGCT 795  
QY 757 cagaagaataatgctgaataatgatgtcttttttcgacgaagcacacaaagaagggaag 816  
DB 796 CAGGAACACACTGGCGCGGAAGTAGCAATCGCATCTCAAGGATGCCGAAGGACCCGGAATG 855  
QY 817 aagaaatcgcaataatgsgtctgttctcaactgagtactttgtcatttaacgtgct 876

DB 856 AAGTCAAAAGTGATCTTGGCCATCATGGTGGGAGCATTTACTATGCATCATGTATCTCAAC 915  
QY 877 atgcactggccttttgaaaaggttttcgcatgtatcagaatggcgaggtt---gcgac 933  
DB 916 TATGGAGTTGGCTTCTGGATGGGGTCTCGGTCTCCTGTCGAAAGAAATCAGTAACATCAAA 975  
QY 934 gttggcaagttcttactgttgctcttcgctacccttagcagcaogtccatctcaatg 993  
DB 976 GCGGGATGTGCTGACCATCATATGATGCCATCATCTAGTTCCTATAATCTGGGCAAT 1035  
QY 994 cttgcgcttcaggttcagtcgttttaacaaacgcgcgcatcttcggctcogaatttcaagt 1053  
DB 1036 GTGCCCCCAATGACAAAGCGCTGTGCGACGCTGTGCGGCGAGCATCAAAAGCTGTATGGC 1095  
QY 1054 atcattgacaaacccacgcagctcgcacctctcgacccttcgaaagcagcagagggc 1113  
DB 1096 ACCATTGACCGTCAATCCCTTTAGATGCGGTTGTCCGATCAGGGCAAGACCTGGAGTTT 1155  
QY 1114 tgctagggtcaaatgttagatccaaacccctggcatttgctccctcccgaccatctgcc 1173  
DB 1156 GTCGAGGAATATTGTACTCCAGAACATCAGGCACGTTTATCTCCTCGACCCGAGGTC 1215  
QY 1174 caagtacttcgagatttcaacttgacaattccagctggcagcagcggccctcgctcgt 1233  
DB 1216 ACTGTTGCCATGATCTGAGTTGCTATATCTCCTGCCGTAAGAGACACTGCTTCGTGGGT 1275  
QY 1234 gcatcagtagcggcaaaaagcacaatggtcgcttacttgaaacggtggtatctgcccagt 1293  
DB 1276 CCGTCAGGGTCCGCAAAAGTACCATCATCAGCTTGCTTGAGCGGTTTATGACCCAGTC 1335  
QY 1294 tcggggaggaattactctgattgggttggaactgggacaatacaatgtgaaatggctgaga 1353  
DB 1336 GCGGGACCATATTGCTCGACGGCCACGATATCCAGACATTTGAATCTCCGATGGGTGCGA 1395  
QY 1354 agccgcatctgcctcgttcaacacaggaacctgtgttcttctgycacaatcttcagaac 1413  
DB 1396 CAGCAAAATGTCGCTCGTTCGCAAGAGCGGAGACTCTTTCGCCACACCATCCGGGAGAT 1455  
QY 1414 attgcaacgggtttcattgtagtagcaacgagatctgctcgcgcaaaaaacaaat---ggag 1470  
DB 1456 ATCGGTTATGTATCTATTTGGCTCTCGATTTCGAAAGGAATTCAGCTAGAGATCCGAAAG 1515  
QY 1471 cttgtcaaaaagcttgcgaagccagcaatggcgactgttcttaattatgacttcgcaac 1530  
DB 1516 CGACTCGAAGCAGCAGCTCGAATGCAAAATGCGCAGCACTTCATCATGGCGCTTCGGAAT 1575  
QY 1531 ggttatgagactgaagtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1590  
DB 1576 GGTATGATACAAATATCG-----AGAGCTTCTCGCTTTCGGGGGCGCAGAGCAGCGC 1629  
QY 1591 attgcaatcgcaagaagtatcatatcggtcccaagatcctgttactcgatgaagctacc 1650  
DB 1630 ATCGCTATTGCGACAGCTATCATAAAGACCCCTAGATCTTACTCTTAGACCAAGCAACA 1689  
QY 1651 agcgccttgaccggaagcgaggaagtggtccaggaggccttgaacggagtggtccaaa 1710  
DB 1690 TCCGCCTTGGACACGAAATCCGAAAGCTCGTTTACAGGAGCGCTTGCACAGGCTTCTAAG 1749  
QY 1711 gaccgcactacttggctcattgcccacaaactagcactgttcaaaagtctggcacaatc 1770  
DB 1750 GGTAGAACGACTATTGTGATCGCTCATCGCTGTCACAAATCCAGAAAGCATACAAATC 1809  
QY 1771 gcaatcttccagggggaaaatcgtcgagcaaggcacaacacacgaattgatcgatc 1830  
DB 1810 ATTGTTCTTGCAATGTCAGATTGTTGAGCAAGGCGCCGATGAGCATTTGATGGATCGA 1869  
QY 1831 ggctgtcattacgcgcactggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1864  
DB 1870 CCGGGGATCTATTGTGATGATGTTGGAAAGCGCACAGAGATTAAAGAAACGGTACTCGAGATAC 1929  
QY 1865 -----acctggggtgacgaacaaacacgaacatgaggaagccctcgac 1908

Db	1930	TCCAAGAGATATTCCAGCAGTATTGACGAACCTTGTCGCCCTAAGCATAAACCAATGACCTTC	1989
Qy	1909	gaaaggcagcagcaggaagctgctggtgaacgcgcgcgcacttgagcgcactcaccaact	1968
Db	1990	TTCTTCGACAAGATTTACCCGGCGATGACGAGTCCGATATTTTACAGCATCCTTCAGTGAT	2049
Qy	1969	gocacatctcaagctgga	1989
Db	2050	GACGCATCGATATTGGATTGCACACTGGGGAGAAACAAGACCAAGTGTGCGAATGTCA	2109
Qy	1990	ctggagaagcgaaggtgcggtcgggaactttgggtaactcgtctcctaaaatgcatocta	2049
Db	2110	CTATCTCATCTCATCGCCGCTCAAGGAAGAGGCATATTCTTTCTGGACTTTGTTCAG	2169
Qy	2050	atcatgtttctacgaacaaaaaaatctctactggtgcttcttgttgcacaaataacaggtt	2109
Db	2170	TTCTTTGCTCTGTTCAACCGTCCGAGTGGCGGTTTATTTACTAGGGCTCTCGCGATCT	2229
Qy	2110	ctgatgatcgcgggccacattctcaggacaaagccctttgttttcgagattgctcactgtc	2169
Db	2230	ATACTGGCAGGGGGATCAACCATCGCAGGCTGTGCTCTTTGGGAAGGCTGTGCACACC	2289
Qy	2170	ttcggagttagtggctg-----cggcacaggaaacgggcagacttttatattctg	2220
Db	2290	CTCAGCCTGCCCCCTTTGAAATATCCGAAGCTCCGGCAGATGCCAAGTCTTGTCGCCTC	2349
Qy	2221	atgttctttgtcgtggcttagaaactagtaggatattcacgatggctggacatgc	2280
Db	2350	ATGTTTCTCATGATTTGGCATTTGTCTGCTGTGTGTATAGTGTGCAGAGGACCTTGTTT	2409
Qy	2281	aacgttatttcacaagttgtcacccactcgtctcaagcgcgaatgttcccaacgagctag	2340
Db	2410	GCATATAGTTCCGAAAGATGCTACTCCGGGCCCGCAGTCAAGCATTTCCGGGTGATACTG	2469
Qy	2341	gatcaagacatcgaactcctcgacatccggagcaaaatttctgggtcctcacatcgcga	2400
Db	2470	CACAGGATATCTCTTTCTGATCAACAGGAAAAACACAGCAGGCGCATTAACAGGCACCT	2529
Qy	2401	ctgtcagctctacccacgcagttgcaagagttgatcagcaaattttctcatttatactc	2460
Db	2530	CTGAGCGGGGACCAAGAACTAACTGGCATTAGTGGCGTCCACATTTGGGGACTATCTTG	2589
Qy	2461	gttg--tcgggtcaacatcgtctcgagcagtgctctaccactagcctatggatggaaactg	2518
Db	2590	ATTGCTCTGTCAATTGTGTGCTTCTTGGGGTAGCCCTGTGTAATAGGCTGGAAGTTG	2649
Qy	2519	ggcctgggttggttggttggtgcacttccaccctcgttttggctgggtacactcagaatt	2578
Db	2650	GCATGTGTTGCATCT---CCGCGGGTTCCAGCCCCGTGATGTGTGGATTTGTTGCGGTT	2706
Qy	2579	cgtctagacgaagctagaagcgcggaactcgcgaactttgcgaactttgcgaagtgctggcct	2638
Db	2707	TGGATGTTGGAGCGGTTCCACGAAGGGCCCAAGAAAGCATATCAAGAATCAGCTAGTTCT	2766
Qy	2639	gcaagcgaagcagttacgcgatccggaccctctcatctttgactctctogaagggccatgtt	2698
Db	2767	GCTTGTGAGGCAGCTCTGCCATCCGTACCGTCTGTTCTTTGACCATGGAGACGGAGCT	2826
Qy	2699	ctccaacagttactcggacatgttgagcaaggtcttgctaaagatactcaaaagcttttggct	2758
Db	2827	TTGCAATCTCATAGCAGACAGCTCCGTGGGCGCTGGAAGATGATATACTCCCCATTTG--	2884
Qy	2759	ttggacgatgttttgggttttcaacttgtaacagtcgatgaggtttttggctattggccctgg	2818
Db	2885	TGAAATCGTCTTACTCTACGCCAGCTCACAGGCGTTGCCCTTCTTCTGCGATGGCGTTGG	2944
Qy	2819	gatttt--gtattgcagtcgataattggtcttcaggtcgagtagcacacaactcaatttata	2877
Db	2945	GGTTCGTGATGGCGGCTCTACTTGGTCTATGAGAGTATTCCTCTGTTTCAATTTTACG	3004
Qy	2878	tcaatctcggggcggttttgggtcgggttccaagcagcagccagatatttggcttaactcc	2937
Db	3005	TGTGCTTCAGTGAGGTGATCTTTTG--GGGCCCAAGCCGCTGGAACGGTTTCTCTCATGCC	3063

QY	2938	acgagttttaccaaggctcggtcggtcgagaaactatctctggtcgggacattgaag	2997
DB	3064	CGGGATATGGGAAAAGCGAAGCACGCTGGCCGTGAGTTCAAACGCTGTTTCA--GCAGC	3120
QY	2998	cggaccatccggaacggagaggaacaagaataaaaggcccgatgggtgatgccctgtc	3057
DB	3121	GACACAATGCATGCTTCTCGGAGCAAGGAGTTCCCGGTTACGTCATGCGAGACATGGTG	3180
QY	3058	gacctcgagagcattgaattcagggtatcgtaacgtgattcggctcgagttctccgcggg	3117
DB	3181	GAGTTCGGCGATGTCCTCTTCGATATCCAGTCGCTTGGAAACAACAATCCTTCGCCAC	3240
QY	3118	gtttccatgacaatcagagccaggaacaatttgtagcttatgtgggcgtctbtggctgtggc	3177
DB	3241	TTGAACCTGACCATCAAGCCAGGGCAATTTGTGTGACTAGTGGTGGAGTGAAGCGGC	3300
QY	3178	aagtcacagttgatcgctttgctgaacgattctacgaccgcgacctcgggcgaattcca	3237
DB	3301	AAGAGCACTACCATTTGGGTGCTGGAGAGATTCTATGATCCCGCTCAAGAGTGGTGTTTAC	3360
QY	3238	tttgcacacggaatatatgcgaataatgtccgcgcgcttgcacgcggccatatgtctttg	3297
DB	3361	GTAGATGAAAGAATAATTATCACTCTCGAGATGTCTTCTATCGCAGTCACTTGGCTCTC	3420
QY	3298	gtccaaacggaaaccacacactttaccaaggctccggttcgcgagaaatgtgacgttggccctc	3357
DB	3421	ATTAGCCCAAGAACCCGACCTTTGTTCCAAGGGAGCATCCGAGAGAACATCTCTGTAGGGAGC	3480
QY	3358	ga-agccgaaattatcagaagaagctttgtcaaggagccttcctccgaaggccaatgtttg	3416
DB	3481	AACACACCTCATGTCCACAGATGATTTCTTGTAAAGCCTGCAAGGATCCAAACATATAC	3540
QY	3417	gattttgtcatctctttaccgaagcgttttgaaacgccttgcggtcacaaggaggaatga	3476
DB	3541	GATTTTATTATTGTCTATTCGGCAAGGCTTCAACAC-CATTGTTGGCAATAAGGGAGGCAT	3599
QY	3477	gtctccgcgcgggcaacacagcggatcgccatcgcaagagcattgattcgaataccaaa	3536
DB	3600	GCTCTCCGGTGGCCAGAACAGCGCATTTGCCATCGCTGGGCCCTCATTTTCGGAAATCCAAA	3659
QY	3537	gctgttgtaattgacgaagcgagctcagccctcgacacgcaatcggaacgtctgttcca	3596
DB	3660	AATCCTCTGTTGGACGAGGCTACGTCGCGCGTAGATTCCGAGTCGGAAAGGTGGTCCA	3719
QY	3597	agctgccctcgataggccatccagagcggaacgacaaatagcagtgggcgcaaccgactttc	3656
DB	3720	GGCAGCATTTGATGCCCGCGCGGGGACGACAAACGATCGCAGTTGCTACCCGGCTGAG	3779
QY	3657	cactattcggaaatgttgatgtattttttgttttgcacacgggaaatcgccgaaacggg	3716
DB	3780	CACATCCAGCGCAGACTTGATATATGTCTAGATCAGGGGAGGTCGTTGAATCAGG	3839
QY	3717	cactcaacgggaactacaacgactgagagggaatatattaca	3758
DB	3840	TACACCCGGGAACCTTCTCCGAAAGAGGGACGCTACTACGA	3881

RESULT	14
FNU62929	
LOCUS	
DEFINITION	FNU62929 Filobasidiella neoformans multidrug resistance protein 1 (mdr1) mRNA linear PLN 09-DEC-1997 mRNA, complete cds.
ACCESSION	U62929
VERSION	U62929.1 GI:2668552
KEYWORDS	.
SOURCE	Filobasidiella neoformans. Filobasidiella neoformans
ORGANISM	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
REFERENCE	1 (bases 1 to 4227)
AUTHORS	Thorneswell,S.J., Peery,R.B. and Skatrud,P.L.









```
Db 4104 TTCTACCGGATTACTTCTCTGAAGGCGAGGTAGCCGAGCATGGAACACATCAGGA 4163
QY 3729 actcaacgactgagagaagatatattacgagatg 3762
Db 4164 GCITTTGCCAAGAAGGGTGGCTATTATGAGCTG 4197

RESULT 15
AF082072 AF082072 4872 bp DNA linear PLN 20-APR-1999
LOCUS Emericella nidulans ABC transporter protein AtrC (atrC) gene,
DEFINITION complete cds.
ACCESSION AF082072
VERSION AF082072.1 GI:4587970
KEYWORDS
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
1 (bases 1 to 4872)
Angermayr, K., Parson, W., Stoffler, G. and Haas, H.
Expression of atrC-encoding a novel member of the ATP binding
cassette transporter family in Aspergillus nidulans-is sensitive to
cycloheximide
Biochim. Biophys. Acta 1453 (2), 304-310 (1999)
99156773
2 (bases 1 to 4872)
Angermayr, K., Parson, W. and Haas, H.
Direct Submission
Submitted (04-AUG-1998) Institut fuer Mikrobiologie Medizinische
Fakultaet, Universitaet Innsbruck, Fritz-Pregl-Str. 3, Innsbruck
A-6020, Austria
FEATURES
Location/Qualifiers
1. 4872
/organism="Emericella nidulans"
/strain="FGSC A4"
/db_xref="taxon:162425"
/chromosome="11"
/map="W21/A03"
join(<524..1539,1596..3583,3650..>4527)
/gene="atrC"
/product="ABC transporter protein AtrC"
<524..>4527
/gene="atrC"
join(524..1539,1596..3583,3650..4527)
/codon_start=1
/product="ABC transporter protein AtrC"
/protein_id="AAD25925.1"
/db_xref="GI:4587971"
/translation="MYRKSDRDNLLEFVHGRYDWRNLVAFMSGLGEKARTVLDVKD
DDPRSGKATPRWNGYSVQDHVIRVTGLFAAIAAGCAALPMWTLVFGTLIDFNWD
GAKLSPDFRSHVSNQALWFTYLFISIFVLSFNTACLRLTATRCVRLRHDFIRSI
LRQDLSFNCPLPGTVATVLSNNADLVEIGLGEKLGIAIEGVAQLCAAFVVAFAROWK
LTIQVADTLAMLVIVTVILETRITTKILATYSKAGGIAEALASTHIVKAYNAAS
KLQARVDTYLERATOLGVKGRPIRGIOYGAQPAIMECAVALAWFYGLLVLKGEIESG
GYLTILTSLVIGSQSLTILGPFIGEVSKTAAQAQELFQVIDRKNPIDSLSGRILT
GVTKLSFRNVSFPRSVRVLDVTVDFEAGTKTAIVGSSGSGKSTILALVSRFF
DPVSGSLVDGHP IHELNIRLWLGQISVQOEPVLFSEIFANVCHGFRTDMLLPE
HERRIVQACEAAFAHFITQGLPEQYDTPVAGDGLLGGQKQRIATARSIIIRNPI
LLIDEATISALDPNAEGTVQAAALNNSKTRTIIIVAHRLSTVORADNIVLVRKRVBO
GSHRELLAKGTVYFDLVAAGTDEGIISTAGDQGTSTFADSVLYKKAEDYNAMSSQ
SKATAHERVPLLKCLFILLRGLRWLPLFFGLVLVLRKROCLPCSSGSRFOOHLFI
FQPLPAMASEMLKGNFWGIIYIVLAVSLVICAGLGFVFFVAASFSLTYSRYPFA
VMNRLOVFEEDQESAGVMTGQSLDQRIEDLISLCIGFILLVNVNVLASILALA
AGRLALVAIFGLPPLFLAGYVVRLEITCOERTRLVLESARFATEAISRTVAS
LYNEVTFVWYVAVIFEGQAAGFLGTYVNTAKHTAANNIIHLGSRPSINASTKQ
EYVQLDSDTALEKDVHFSYPARTVPVVLQGLSFVKKGEHGLVAGSGCKTIVISL
LERFYAGSGEITFINGIPLHDDIVHSHRARIQVLTQPTLYQGSIRDNVLIGISHQ
NEELDQSIIEKTKAYKDANIODFTQSLPEQQTDPGTRGLALSGQRQRIATARALI
RDELPLLFDEATSAIDTENERLVQEAIERVSHGPGRTTISVAHRLTVRRCDRLVLH
EGRVEEGTHAELMARGGRYQWVLAQLDR"
```

```
BASE COUNT 1159 a 1272 c 1250 g 1191 t
ORIGIN
Query Match 9.0%; Score 343; DB 8; Length 4872;
Best Local Similarity 46.2%; Pred. No. 7.5e-85;
Matches 1778; Conservative 1; Mismatches 1901; Indels 165; Gaps 13;
QY 96 agactatctacgaatcttcagatatccgacaaaacagcactggactctcaatgtcatcgc 155
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 AGCTTATTTTCGAGTCTGGGGTACTCGAGCGGCCAGGATCATGTGATCAGGCTCACTGG 744
QY 156 gtcacatctgcgcacatcgagatccggggtctccctctctctctctctctctctcgttag 215
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 ACTCTTCGACGCTATTGTCATGCGGTGACGCTTTTACCAGATGATGACTCTTGTGTGCGGTAC 804
QY 216 cttcaccaagaattcaacaattacaattcgggcgcgggagtcctcgaaagcgtcaagggc 275
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 805 TCTATTCGACAACTTCAACGACTGGGGCGCGGGAAGCTCTCGCCTGATGATGAGTTCGGGTC 864
QY 276 cgatgtggatcttctgctcctgtggtctctacctcttcttatttggaagtttgcctcac 335
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 865 TCACGCTCTCAGAAATGCCTTGTGTTACAGTATCTCTTCATTAGCATCTTGTGTCTCTC 924
QY 336 gtaegtcttcacggctgcattaccatttcagatatcaagaaccactcgaactcttcgacg 395
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 GTCAATTCAACACTGCGTGTCTAGACTGACCGCTACCCGGTGCCTCAGGCGCCTCAGACA 984
QY 396 agtgttctctgaatgcacattcggaaggtctctgacatttcgacaaagcagagaatgg 455
Db 985 CGACTTCATCCGTTCAATCCTCGCCAGGATCTTTCCTACTTCGACAAATTCGCTGCCAGG 1044
QY 456 agcaatcgccactcargtcactaccatggaacccgtatatacaaacaggattatgccagaa 515
Db 1045 CACGGTGGCAACCGTGTCTATCCAACAATCAGAGACCTTGTGAGATCGGTCTCGCGAGAA 1104
QY 516 attggtctttacgtgcgaggaactttcaatgtttcttcttcttcttcttcttcttcttctt 575
Db 1105 GCTTGGAAATCGCAATTGAGGGCGTGGCACAGCTTTCCGCTGCAATTTGTCTGCCCTTCGC 1164
QY 576 gtctcagtggagactagcttttaacacccatgtctcgtcactccctgccatttctctgtgcc 635
Db 1165 TCGGCAGTGGAAATTAACCTCTGGTCTGCTCGTCTCGGACTCTTCCACTCGCGATGTTGTGAT 1224
QY 636 cggcactctgcacgaatgtatgcggtcagaggccagagatcaccagatctactcacg 695
Db 1225 TGTCTGCTACTGTAATCTTAGAGACAAGGATCACTACAAAGATCCTGCCCATTTACAGTAA 1284
QY 696 cgcgcgtctcctcgagaagaagtattatcatccatcccgagacagtcacgtttcttctacgc 755
Db 1285 GGCCCGGGGATCGCGGAGGAGCGGTGGCCAGTACTATATGCTCAAGCGGTACAAACGC 1344
QY 756 ccagaagaaaaaggtcgaaaaataatgatgtctttttgcagcaagcacacaagaagggaa 815
Db 1345 TGCTTCAAAGTTGCAAGCTAGGTACGACTCGTACTTGGAGAGGGCGACGACGCTCGAGT 1404
QY 816 gaagaaatcgccaaataatgsggtctgttctcaactgagtagtacttttgcatttaccgtgc 875
Db 1405 CAAAGAGGGCGCGATCCGCGGTATACAGTACGCGGGTCAATTCGCAATCATGTTTTCGCG 1464
QY 876 tatcgactgcgcttttggaaaggttttcgactgtatcaagaatgcgaggttgcgcacgt 935
Db 1465 CTACGCGCTCGCTTGGTTCTACGGAATTCGACTCCTTGTGAAAGGCGAGATCGAAAGCGG 1524
QY 936 tggcaagctcttact----- 951
Db 1525 CGGCTACCTGATCAGCTAGGAGCGCTGCCCTGAACCGCGTTTAGACAAATATAGTGCTTA 1584
QY 952 -----gttgtcctttccgtcaacttagcagccacgtccatctcaatgttgcg 999
Db 1585 ACTGACATAAAGAGTCTTACCCTACCTAGTCTCTGATAGGAGGCAATCCCTCACGCTTATTGGC 1644
QY 1000 ctttcaggttcagttcggtttaccacgcgcgcacatctctcgttcgcgaattatttcagtatcatt 1059
```









Query Match 0.6%; Score 21; DB 2; Length 4002;  
Best Local Similarity 100.0%; Pred. No. 0.8;  
Matches 21; Conservative 0; Mismatches 0; Indels

Qy 3553 gaagcgagctcagccctcgac 3573  
|||||  
Db 3763 GAAGCGAGCTCAGCCCTCGAC 3783

## RESULT

```

US-08-996-545-3
: Sequence 3, Application US/08996545
: Patent No. 5928898
:
: GENERAL INFORMATION:
:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Waard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene at rD of
: TITLE OF INVENTION: Aspergillus nidulans
:
: NUMBER OF SEQUENCES: 3
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
:
: ZIP: 46285
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/996-545

```

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4002 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-996-545-3

```

Query Match 0.6%; Score 21; DB 2; Length 4002;  
Best Local Similarity 90.5%; Pred. No. 0.8;  
Matches 19; Conservative 2; Mismatches 0; Indels

Qy	3553	gaagcgacgtcagccctcgac	3573
		:     :	
Db	3763	GAAGCGAGCUCAGCCCUCCGAC	3783

### RESULT

US-09-328-320-1  
; Sequence 1, Application US/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.

```

: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
: TITLE OF INVENTION: Aspergillus nidulans
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/328,320
: FILING DATE:
:

```

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/996,545  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763

```

1: INFORMATION FOR SEQ ID NO: 1:
2:
3: SEQUENCE CHARACTERISTICS:
4: LENGTH: 4002 base pairs
5: TYPE: nucleic acid
6: STRANDEDNESS: single
7: TOPOLOGY: linear
8:
9: MOLECULE TYPE: cDNA
10:
11: HYPOTHETICAL: NO
12:
13: ANTI-SENSE: NO
14:
15: FEATURE:
16:
17: NAME/KEY: CDS
18:
19: LOCATION: 1..4002
20:
21: US-09-328-320-1

```

Query Match 0.6%; Score 21; DB 4; Length 4002;  
Best Local Similarity 100.0%; pred. No. 0.8;  
Matches 21; Conservative 0; Mismatches 0; Indels

Qy 3553 gaagcgacgtcagccctcgac 3573  
|||||

Db 3763 GAAGCGACGTcAGCCCTCGAC 3783

## RESULTS

```

RES001
US-09-328-320-3      Application US/09328320
: Sequence 3
: Patent No. 6228615
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Waard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug
: TITLE OF INVENTION: Aspergillus
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285

```

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/328,320  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4002 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-09-328-320-3

Query Match 0.5%; Score 21; DB 4; Length 4002;  
Best Local Similarity 90.5%; Pred. No. 0.8;  
Matches 19; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3553 gaagcgagctcagccctcgac 3573  
|||||:|||||:|||||  
Db 3763 GAAGCGAGCUCAGCCUCCGAC 3783

RESULT 5  
US-08-479-744A-32/c  
; Sequence 32, Application US/08479744A  
; Patent No. 6084067  
; GENERAL INFORMATION:  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; APPLICANT: Gray, Gary S.  
; TITLE OF INVENTION: No. 6084067el CTLA4/CD28 Ligands and  
; NUMBER OF INVENTION: Uses Therefor  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/479,744A  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/280,757  
; FILING DATE: 26-JUL-1994  
; APPLICATION NUMBER: 08/109,393  
; FILING DATE: 28-AUG-1993  
; APPLICATION NUMBER: 08/101,624

; FILING DATE: 26-JULY-1993  
; APPLICATION NUMBER: 08/147,773  
; FILING DATE: 3-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: RPI-004CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 39 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: oligonucleotide  
; US-08-479-744A-32

Query Match 0.5%; Score 18; DB 3; Length 39;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1974 atctcaagctggagacct 1991  
|||||:|||||:|||||  
Db 24 ATCTCAAGCTGGAGACCT 7

RESULT 6  
US-08-280-757B-32/c  
; Sequence 32, Application US/08280757B  
; Patent No. 6130316  
; GENERAL INFORMATION:  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; APPLICANT: Gray, Gary S.  
; APPLICANT: Greenfield, Edward  
; TITLE OF INVENTION: No. 6130316el CTLA4/CD28 Ligands and  
; NUMBER OF INVENTION: Uses therefor  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,757B  
; FILING DATE: 26-JUL-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/101,624  
; FILING DATE: 26-JULY-1993  
; APPLICATION NUMBER: 08/109,393  
; FILING DATE: 19-AUG-1993  
; APPLICATION NUMBER: 08/147,773  
; FILING DATE: 3-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: RPI-004CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 39 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: oligonucleotide  
US-08-280-757B-32

Query Match 0.5%; Score 18; DB 3; Length 39;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1974 atctcaagctggagacct 1991  
Db 24 ATCTCAAGCTGGAGACCT 7  
|||||

RESULT 7  
US-08-740-309-1  
; Sequence 1, Application US/08740309  
; Patent No. 5919623  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: NUCLEIC ACID ASSAYS  
; NUMBER OF SEQUENCES: 3  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/740.309  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9408344.1  
; FILING DATE: 27-APR-1995  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 613 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-740-309-1

Query Match 0.5%; Score 18; DB 2; Length 613;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1568 tgagtggagggtcaacgac 1585  
Db 353 TGAGTGGAGGTCAACGAC 370  
|||||

RESULT 8  
US-08-286-889-45  
; Sequence 45, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: P38,659  
; REFERENCE/DOCKET NUMBER: 27866/32168  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 52..3519  
US-08-286-889-45

Query Match 0.5%; Score 18; DB 1; Length 3519;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 879 cgcactggccttttgaa 896  
Db 1281 CGCACTGGCCTTTTGAA 1298  
|||||

RESULT 9  
US-08-485-618-45  
; Sequence 45, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,618  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652



;  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32797  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 52..3519  
US-08-485-618-45

Query Match 0.5%; Score 18; DB 1; Length 3519;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 cgcactggccttttgaa 896  
|||||

Db 1281 CGCCTGGCCTTTGGAA 1298

RESULT 10  
US-08-362-652-45  
; Sequence 45, Application US/08362652  
; Patent No. 5766850  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,652  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 52..3519  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3519 base pairs

;  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 52..3519  
US-08-362-652-45

Query Match 0.5%; Score 18; DB 1; Length 3519;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 cgcactggccttttgaa 896  
|||||

Db 1281 CGCCTGGCCTTTGGAA 1298

RESULT 11  
US-08-605-672-45  
; Sequence 45, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605,672  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 52..3519  
US-08-605-672-45

```
Query Match          0.5%; Score 18; DB 1; Length 3519;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 879 cgactggccttttgga 896  
 |||||  
 Db 1281 CGCACTGGCCCTTTGGAA 1298

```

RESULT 12
US-08-482-293A-45
; Sequence 45, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A

```

FILING DATE: 530  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/173,497  
 FILING DATE: 23-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/286,889  
 FILING DATE: 5-AUG-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/362,652  
 FILING DATE: 21-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Williams Jr., Joseph A.  
 REGISTRATION NUMBER: 38,659  
 REFERENCE/DOCKET NUMBER: 27866/32684  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-474-6300  
 TELEFAX: 312-474-0448  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 45:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3519 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:

Query Match 0.5%; Score 18; DB 2; Length 3519;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 879 cgcactggccttttgaa 896  
 |||||  
 pb 1281 CGCACTGGCCTTTTGGAA 1298

RESULT 13  
US-08-943-363-45  
; Sequence 45, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/943,363  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 52..3519  
; US-08-943-363-45

```
Query Match          0.5%; Score 18; DB 2; Length 3519;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 879 cgcactggccttttgga 896  
 |||||  
 Db 1281 CGCACTGGCCTTTTGGAA 1298

```

RESULT 14
US 09-193-043-45
: Sequence 45 Application US/09193043
: Patent No. 6251195
: GENERAL INFORMATION:
: APPLICANT: Gallatin, Michael W.
: APPLICANT: Van der Vieren, Monica
: TITLE OF INVENTION No. 6251395el Human 2

```

FILE REFERENCE: 27866/35004  
CURRENT APPLICATION NUMBER: US/09/193,043  
CURRENT FILING DATE: 1998-11-16  
EARLIER APPLICATION NUMBER: 08/173,497  
EARLIER FILING DATE: 1993-12-23  
EARLIER APPLICATION NUMBER: 08/286,889  
EARLIER FILING DATE: 1994-08-05  
EARLIER APPLICATION NUMBER: 08/362,652  
EARLIER FILING DATE: 1994-12-21  
EARLIER APPLICATION NUMBER: 08/943,363  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 45  
LENGTH: 3519  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (52)..(3516)  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: primer  
US-09-193-043-45

Query Match 0.5%; Score 18; DB 4; Length 3519;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 cgcactggccttttggaa 896  
|||||  
Db 1281 cgcactggccttttggaa 1298

## RESULT 15

US-08-485-618-52  
Sequence 52, Application US/08485618  
Patent No. 5728533  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32797  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3803 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3486  
US-08-485-618-52

Query Match 0.5%; Score 18; DB 1; Length 3803;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 cgcactggccttttggaa 896  
|||||  
Db 1230 cgcactggccttttggaa 1247

Search completed: April 27, 2002, 05:56:08  
Job time: 15455 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 02:58:50 ; Search time 4999.95 seconds  
(without alignments)  
10236.193 Million cell updates/sec

Title: US-09-882-694A-10  
Perfect score: 3792  
Sequence: 1 atggcagatgaatcgagaa.....aatcttagaccaagcatga 3792

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	22	0.6	468	10	BM137266	BM137266 WHE2631_F
C 2	21	0.6	325	9	AI594082	AI594082 vo52c01.y
C 3	21	0.6	327	9	AA608457	AA608457 vo52c01.r
C 4	21	0.6	334	9	BB670692	BB670692 BB670692
C 5	21	0.6	343	9	BB670723	BB670723 BB670723
C 6	21	0.6	359	9	BB803996	BB803996 BB803996
C 7	21	0.6	385	9	AU020440	AU020440 AU020440
C 8	21	0.6	399	9	BE113735	BE113735 UI-R-BJ1-
C 9	21	0.6	399	10	BM213208	BM213208 C0833802-
C 10	21	0.6	402	9	BB735228	BB735228 BB735228
C 11	21	0.6	403	9	AI549109	AI549109 UI-R-C3-t
C 12	21	0.6	404	9	BB672487	BB672487 BB672487
C 13	21	0.6	404	9	BB816748	BB816748 BB816748
C 14	21	0.6	430	9	AI643840	AI643840 vu02g12.x
C 15	21	0.6	432	9	BB670914	BB670914 BB670914
C 16	21	0.6	435	10	BG381149	BG381149 UI-R-CS0-
C 17	21	0.6	450	9	BB729150	BB729150 BB729150

18	21	0.6	454	10	BF012195	BF012195 us39d07.y
C 19	21	0.6	457	9	AI713238	AI713238 UI-R-Y0-a
C 20	21	0.6	469	10	BM214273	BM214273 C0848C04-
C 21	21	0.6	474	9	AI050414	AI050414 ub30f07.f
C 22	21	0.6	495	10	BM211478	BM211478 C0807B12-
C 23	21	0.6	501	9	AI021014	AI021014 us98f11.f
C 24	21	0.6	511	10	BM211846	BM211846 C0813C07-
C 25	21	0.6	513	9	AI987703	AI987703 um05f12.x
C 26	21	0.6	514	10	BM216258	BM216258 C0877G09-
C 27	21	0.6	517	10	BF543407	BF543407 UI-R-C3-t
C 28	21	0.6	518	9	BE101140	BE101140 UI-R-BJ1-
C 29	21	0.6	533	9	AI155155	AI155155 ud80h10.f
C 30	21	0.6	541	10	BM214673	BM214673 C0854C01-
C 31	21	0.6	545	10	BM217358	BM217358 C0894G05-
C 32	21	0.6	545	10	BM217367	BM217367 C0894H03-
C 33	21	0.6	562	10	BM217121	BM217121 C0891C06-
C 34	21	0.6	565	10	BM214433	BM214433 C0850F12-
C 35	21	0.6	571	10	BM215931	BM215931 C0872H08-
C 36	21	0.6	576	10	BM215346	BM215346 C0864B08-
C 37	21	0.6	585	10	BM211882	BM211882 C0813H09-
C 38	21	0.6	590	10	BM215513	BM215513 C0866G10-
C 39	21	0.6	592	10	BM212432	BM212432 C0821G10-
C 40	21	0.6	617	10	BM213756	BM213756 C0841C01-
C 41	21	0.6	620	10	BM214102	BM214102 C0846B05-
C 42	21	0.6	638	9	BB795396	BB795396 BB795396
C 43	21	0.6	644	10	BM213718	BM213718 C0840E05-
C 44	21	0.6	644	10	BF147586	BF147586 us39d07.x
C 45	21	0.6	661	12	AZ864595	AZ864595 ZM0174G13

ALIGNMENTS

BM137266 468 bp mRNA linear EST 28-NOV-2001  
WHE2631\_F08\_L152S Wheat Fusarium graminearum infected spike cDNA  
library Triticum aestivum cDNA clone WHE2631\_F08\_L15, mRNA  
sequence.

BM137266  
BM137266.1 GI:17146033  
EST.  
Triticum aestivum  
bread wheat.

ACCESSION

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 468)  
Anderson, O.D., Chao, S., Han, P.S., Heinlen, S., Hsia, C.C., Kang, Y.,  
Kruger, W.M., Lazo, G.R., Miller, S., Muehlbauer, G.J., Miller, R.,  
Pritsch, C., Rausch, C.J., Seaton, C.L., Tong, J.C., Vance, C. and  
Wilson, C.F.

TITLE

The structure and function of the expressed portion of the wheat  
genomes - Fusarium graminearum infected spike cDNA library

JOURNAL

Unpublished (2001)  
Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA

COMMENT

Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov  
Sequences have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20. No effort was taken  
to identify ESTs of fungal origin from this library, thus this EST  
could be of wheat or fungal origin.  
Seq primer: Stratagene SK primer.

FEATURES

source

1..468  
/organism="Triticum aestivum"  
/cultivar="Sumai3"  
/db\_xref="taxon:4565"  
/clone="WHE2631\_F08\_L15"

```

/clone_lib="wheat Fusarium graminearum infected spike cdna
library"
/tissue_type="Spike"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/Note="Vector: Lambda Uni-ZAP XR, excised phagemid
pluascript SK; Site_1: EcoRI; Site_2: XhoI; Plants were
grown in the greenhouse. Spikes were sprayed at anthesis
with Fusarium graminearum. Total RNA, and poly(A) RNA were
prepared and pooled from infected spike at 0, 6, 12, 24,
36 and 48 hours after inoculation, a cDNA library was made
, and the cDNA clones were in vivo excised to give
pluascript phagemids in G. Muehlbauer lab at the
University of Minnesota (Kruger, W.M., Muehlbauer, G.J.,
Pritsch, C., Vance, C.). The cDNA library should contain
genes of both wheat and fungal pathogen origin. Plasmid
DNA preparations and DNA sequencing were performed in the
OD Anderson lab (all other authors)."
BASE COUNT      100 a 183 c 116 g 68 t 1 others
ORIGIN

Query Match      0.6%; Score 22; DB 10; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.4; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 799 gcacaccaagaagggaagaaga 820
|||||
DB 447 GCACACCACAGGAAGAAGA 426

RESULT 2
AI594082
LOCUS
DEFINITION
vo52c01.y1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
clone IMAGE:1033504 5' similar to gb:S60080 DNA POLYMERASE EPSILON,
CATALYTIC SUBUNIT A (HUMAN);, mRNA sequence.
ACCESSION
AI594082
VERSION
AI594082.1 GI:4603130
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 325)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,K., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:585080
This read is a RESSEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .325
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1033504"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"

FEATURES
source

```

```

/lab_host="DH10B"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGGATCCTTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library constructed by Bob Barstead. *
BASE COUNT      104 a 87 c 64 g 70 t
ORIGIN

Query Match      0.6%; Score 21; DB 9; Length 325;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 aatcttcagaacattgcca 1421
|||||
DB 26 AATCTTCAGACATTCGCA 46

RESULT 3
AA608457
LOCUS
DEFINITION
vo52c01.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
clone IMAGE:1033504 5' similar to gb:S60080 DNA POLYMERASE EPSILON,
CATALYTIC SUBUNIT A (HUMAN);, mRNA sequence.
ACCESSION
AA608457
VERSION
AA608457.1 GI:2455314
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 327)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:585080
Pucative full length read
vector to vector length is 328
Seq primer: -28mi3 rev2 ET from Amersham.
Location/Qualifiers
1. .327
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1033504"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT

```



genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi.K., Fujiwake.S., Inoue.K., Togawa.Y., Izawa.M., Ohara.E.,  
Watahiki.M., Yoneda.Y., Ishikawa.T., Ozawa.K., Tanaka.T., Matsura  
.S., Kawai.J., Okazaki.Y., Muramatsu.M., Inoue.Y., Kira.A. and  
Hayashizaki.Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno.H., Fukunishi.Y., Shibata.K., Itoh.M., Carninci.P., Sugahara  
.Y. and Hayashizaki.Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.

```

BASE COUNT      107 a    97 c    62 g    77 t
ORIGIN

Query Match      0.6%; Score 21; DB 9; Length 343;
Best Local Similarity 100.0%;
Matches 21; Conservative 0; Mismatches 0; Gaps 0;

```

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1978, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Research	1985, Vol. 88, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 4, pp. 1-10
5. The Role of the School in the Community	Journal of Educational Research	1995, Vol. 98, No. 5, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2000, Vol. 103, No. 6, pp. 1-10
7. The Role of the Teacher in the Classroom	Journal of Educational Research	2005, Vol. 108, No. 7, pp. 1-10
8. The Impact of Technology on Education	Journal of Educational Technology	2010, Vol. 3, No. 8, pp. 1-10
9. The Importance of Parental Involvement	Journal of Educational Research	2015, Vol. 118, No. 9, pp. 1-10
10. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	2020, Vol. 123, No. 10, pp. 1-10

Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suchiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, F., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.



**AUTHORS** Ko, M.S.H., Kitchen, J.R., Wang, X., Threat, T.A., Sun, T., DePalma, G.E.,  
**TITLE** Liang, Y., Kargul, G.J., Sharara, R. and Doi, H.  
 Systematic analyses of genes expressed in eight-cell stage mouse  
 embryos (The ERATO/Doi Project at Wayne State University) (Ko  
 , M.S.H. et al.)  
**JOURNAL** Unpublished (1998)  
**COMMENT** Contact: Hirofumi Doi  
 Doi Biosymetry Project, ERATO  
 Japan Science and Technology Corporation (JST)  
 WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan  
 Email: hdo@boa.jst.go.jp.

**FEATURES** Location/Qualifiers  
 source  
 1..386  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="J0533A09"  
 /dev\_stage="eight-cell stage embryo"  
 /dev\_stage="eight-cell stage embryo"

**BASE COUNT** 84 a 80 c 107 g 115 t  
**ORIGIN**

Query Match 0.6%; Score 21; DB 9; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1401 aatcttcagaaacattgccaa 1421  
 ||||||||||||||||||  
 Db 282 AATCTTCAGAACATTCGCCAA 262

**RESULT** 8  
**LOCUS** BE113735/c  
**DEFINITION** UI-R-BJ1-aws-g-12-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone  
 UI-R-BJ1-aws-g-12-0-UI 3', mRNA sequence.  
**ACCESSION** BE113735  
**VERSION** BE113735.1 GI:8505840  
**KEYWORDS** EST.  
**SOURCE** Norway rat.  
**ORGANISM** Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

**REFERENCE** 1 (bases 1 to 399)  
**AUTHORS** Bernaldo, M.F., Lennon, G. and Soares, M.B.  
**TITLE** Normalization and subtraction: two approaches to facilitate gene  
 discovery  
**JOURNAL** Genome Res. 6 (9), 791-806 (1996)  
**MEDLINE** 97044477  
**COMMENT** Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized atrium at 16.5 dpc library cDNA Library Preparation:  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=Yes.

**FEATURES** Location/Qualifiers  
 source  
 1..399  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"

**AUTHORS** Ko, M.S.H., Kitchen, J.R., Wang, X., Threat, T.A., Sun, T., DePalma, G.E.,  
**TITLE** Liang, Y., Kargul, G.J., Sharara, R. and Doi, H.  
 Systematic analyses of genes expressed in eight-cell stage mouse  
 embryos (The ERATO/Doi Project at Wayne State University) (Ko  
 , M.S.H. et al.)  
**JOURNAL** Unpublished (1998)  
**COMMENT** Contact: Hirofumi Doi  
 Doi Biosymetry Project, ERATO  
 Japan Science and Technology Corporation (JST)  
 WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan  
 Email: hdo@boa.jst.go.jp.

**FEATURES** Location/Qualifiers  
 source  
 1..386  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="J0533A09"  
 /dev\_stage="eight-cell stage embryo"  
 /dev\_stage="eight-cell stage embryo"

**BASE COUNT** 84 a 80 c 107 g 115 t  
**ORIGIN**

Query Match 0.6%; Score 21; DB 9; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1401 aatcttcagaaacattgccaa 1421  
 ||||||||||||||||||  
 Db 282 AATCTTCAGAACATTCGCCAA 262

**RESULT** 9  
**LOCUS** BM213208/c  
**DEFINITION** BM213208-3 NIA Mouse Blastocyst cDNA Library (Long) Mus musculus  
 cDNA clone C0833B02 3', mRNA sequence.  
**ACCESSION** BM213208  
**VERSION** BM213208.1 GI:17771026  
**KEYWORDS** EST.  
**SOURCE** house mouse.  
**ORGANISM** Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Mus.

**REFERENCE** 1 (bases 1 to 399)  
**AUTHORS** Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Tanaka, T., Luo, A. and  
 KO, M.S.H.  
**TITLE** Systematic Analyses of NIA Mouse Blastocyst cDNA Library (Long)  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: C0833 row: B column: 02  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 399  
 POLYA=Yes.

**FEATURES** Location/Qualifiers  
 source  
 1..399  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:C0833B02-3"  
 /db\_xref="taxon:10090"  
 /clone="C0833B02"  
 /clone\_lib="NIA Mouse Blastocyst cDNA Library (Long)"  
 /tissue\_type="Blastocyst"  
 /dev\_stage="3.5-dpc"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
 NotI; Mouse cDNA project by the Laboratory of Genetics,  
 National Institute on Aging (NIA), Intramural Research  
 Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
 a long-transcript enriched cDNA library (Ref. Genome Res.  
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

/clone="UI-R-BJ1-aws-g-12-0-UI"  
 /clone\_lib="UI-R-BJ1"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-BJ1  
 library is a subtracted library derived from the following  
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV  
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,  
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.  
 For a detailed description of the library from which this  
 clone was derived, please visit our web site at  
 ratest.eng.uiowa.edu. The subtraction has been previously  
 described in (Bonaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)  
 TAG\_LIB="UI-R-BJ1"  
 TAG\_TISSUE="atrium at 16.5 dpc"  
 TAG\_SEQ="GATTC"  
 BASE COUNT 94 a 74 c 95 g 136 t  
**ORIGIN**

Query Match 0.6%; Score 21; DB 9; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1401 aatcttcagaaacattgccaa 1421  
 ||||||||||||||||||  
 Db 311 AATCTTCAGAACATTCGCCAA 291

**RESULT** 9  
**LOCUS** BM213208/c  
**DEFINITION** BM213208-3 NIA Mouse Blastocyst cDNA Library (Long) Mus musculus  
 cDNA clone C0833B02 3', mRNA sequence.  
**ACCESSION** BM213208  
**VERSION** BM213208.1 GI:17771026  
**KEYWORDS** EST.  
**SOURCE** house mouse.  
**ORGANISM** Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Mus.

**REFERENCE** 1 (bases 1 to 399)  
**AUTHORS** Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Tanaka, T., Luo, A. and  
 KO, M.S.H.  
**TITLE** Systematic Analyses of NIA Mouse Blastocyst cDNA Library (Long)  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: C0833 row: B column: 02  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 399  
 POLYA=Yes.

**FEATURES** Location/Qualifiers  
 source  
 1..399  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:C0833B02-3"  
 /db\_xref="taxon:10090"  
 /clone="C0833B02"  
 /clone\_lib="NIA Mouse Blastocyst cDNA Library (Long)"  
 /tissue\_type="Blastocyst"  
 /dev\_stage="3.5-dpc"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
 NotI; Mouse cDNA project by the Laboratory of Genetics,  
 National Institute on Aging (NIA), Intramural Research  
 Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
 a long-transcript enriched cDNA library (Ref. Genome Res.  
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

extracted from a pool of 20 Blastocysts. Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen]:

5'-pcactagctctagatcgcgagcgccgctttttttttt-3' from 0.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lona-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NIA).

BASE COUNT 89 a 85 c 107 g 118 t  
ORIGIN

Query Match 0.6%; Score 21; DB 10; Length 399;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 aatcttcagaacattgccaa 1421  
|||||  
Db 302 AATCTTCCAGACATTGCCAA 282

RESULT 10  
BB735228

LOCUS BB735228 RIKEN full-length enriched, 6 days neonate spleen Mus  
DEFINITION cDNA clone F420015G18 3', mRNA sequence.  
BB735228  
VERSION BB735228.1 GI:16134378

KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 402)  
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akaira,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

Location/Qualifiers  
source 1. 402

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="F420015G18"

/clone\_lib="RIKEN full-length enriched, 6 days neonate spleen"

/tissue\_type="spleen"

/dev\_stage="6 days neonate"

BASE COUNT 119 a 111 c 84 g 88 t

ORIGIN

Query Match 0.6%; Score 21; DB 9; Length 402;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 aatcttcagaacattgccaa 1421

|||||

Db 101 AATCTTCCAGACATTGCCAA 121

RESULT 11  
AI549109/c

LOCUS AI549109 403 bp mRNA linear EST 03-JUL-1999  
DEFINITION UI-R-C3-tq-f-08-0-UI-sl UI-R-C3 Rattus norvegicus cDNA clone  
UI-R-C3-tq-f-08-0-UI 3', mRNA sequence.

AI549109

VERSION AI549109.1 GI:4466597

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 403)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:

clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at

LLNL (infoimage.llnl.gov). IMAGE ID=1785641

Seq primer: M3 Forward

POLYA=No.

FEATURES

source

Location/Qualifiers

1. 403

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-C3-tg-f-08-0-UI"  
/clone\_lib="UI-R-C3"

/dev\_stage="adult"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C3 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day embryos, after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dr track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C3) was constructed as follows: PCR amplified cDNA inserts from UI-R-C2p clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C2p library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C3 library. This procedure has been previously described (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996)"

BASE COUNT 96 a 76 c 94 g 137 t  
ORIGIN

Query Match 0.6%; Score 21; DB 9; Length 403;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 aatcttcagaacattgccaa 1421  
|||||  
Db 311 AATCTCCAGAACATTGCCAA 291

## RESULT 12

BB672487 404 bp mRNA linear EST 05-OCT-2001  
LOCUS BB672487 RIKEN full-length enriched, adult male brain Mus musculus  
DEFINITION cDNA clone 3526405D03 3', mRNA sequence.

ACCESSION BB672487  
VERSION BB672487.1 GI:15971708  
KEYWORDS EST.  
SOURCE house mouse.

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS 1 (bases 1 to 404)  
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (AKIMURA,T., et al. 2001)

## TITLE

Unpublished (2001)  
Contact: Yoshihide Hayashizaki

## JOURNAL

COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES  
source

Location/Qualifiers  
1..404  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="RIKEN full-length enriched, adult male brain"  
/sex="male"  
/tissue\_type="brain"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
GAGAGAGAGAGATCCAGTCTTTTATTAATTAATTCACCCCCCCCCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGAGATTCGAGTTAATTAATTAATTCACCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda  
FLC I"  
BASE COUNT 117 a 113 c 85 g 89 t  
ORIGIN

Query Match 0.6%; Score 21; DB 9; Length 404;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 aatcttcagaacattgccaa 1421  
|||||  
Db 123 AATCTCCAGAACATTGCCAA 143

## RESULT 13

BB816748 404 bp mRNA linear EST 19-NOV-2001  
LOCUS BB816748 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus  
DEFINITION musculus cDNA clone G730036J07 3', mRNA sequence.

ACCESSION BB816748  
VERSION BB816748.1 GI:16989377  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 404)  
**REFERENCE**  
**AUTHORS** Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
**TITLE** RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalizaton and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.  
e mouse tissues.  
**FEATURES**  
**source** Location/Qualifiers  
1. .404  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="G730036J07"  
/clone\_lib="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"  
/tissue\_type="lung"  
/cell\_line="RCB-0558 LLC"  
119 a 109 c 82 g 94 t  
**BASE COUNT** 119 a 109 c 82 g 94 t  
**ORIGIN**  
Query Match 0.6%; Score 21; DB 9; Length 404;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1401 aatcttcagaacattgccaa 1421  
|||||  
Db 87 AATCTCCAGAACATTGCCAA 107  
**RESULT** 14  
**LOCUS** A1643840  
**DEFINITION** v002912.x1 Soares\_mammary\_gland\_NBMG Mus musculus cDNA clone IMAGE:1179526 3' similar to SW:DPOE\_HUMAN Q07864 DNA POLYMERASE EPSILON, CATALYTIC SUBUNIT A ;, mRNA sequence.  
**ACCESSION** A1643840  
**VERSION** A1643840.1 GI:4722315  
**KEYWORDS** EST.

house mouse.  
**SOURCE** Mus musculus  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 430)  
**REFERENCE**  
**AUTHORS** NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
**JOURNAL** Unpublished (1997)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: ccapbs-remail.nih.gov  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:637374  
This clone was previously sequenced on the 5' end only, this new data is from the 3' end  
Possible reversed clone; similarity on wrong strand  
High quality sequence stop: 296.  
**FEATURES**  
**source** Location/Qualifiers  
1. .430  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1179526"  
/clone\_lib="Soares\_mammary\_gland\_NBMG"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/notes="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGAATGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."  
95 a 87 c 116 g 130 t 2 others  
**BASE COUNT** 95 a 87 c 116 g 130 t 2 others  
**ORIGIN**  
Query Match 0.6%; Score 21; DB 9; Length 430;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1401 aatcttcagaacattgccaa 1421  
|||||  
Db 305 AATCTCCAGAACATTGCCAA 285  
**RESULT** 15  
**LOCUS** BB670914  
**DEFINITION** BB670914 RIKEN full-length enriched, 10 days embryo Mus musculus cDNA clone 3425403B19 3', mRNA sequence.  
**ACCESSION** BB670914  
**VERSION** BB670914.1 GI:15970135  
**KEYWORDS** EST.  
**SOURCE** house mouse.  
**ORGANISM** Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 432)  
**REFERENCE**  
**AUTHORS** Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.

Search completed: April 27, 2002, 02:59:00  
Job time: 13452 sec

Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.go.jp,  
URL: <http://genome-gsc.riken.go.jp/>  
Carrincci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10): 1617-1630 (2000)  
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Itoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
10 (11): 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carrincci, P., Sugahara,  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. *Genome Res.* 11 (2): 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.

```

BASE COUNT      125 a      121 c      90 g      96 t
ORIGIN

Query Match      0.6%; Score 21; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1401 aatcttccagaacatigccaa 1421
      |||||||
Db 131 AATCTTCCAGAAATTCGCAA 151

```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 07:44:48 ; Search time 95.74 Seconds  
(without alignments)  
7275.887 Million cell updates/sec

Title: US-09-882-694A-10

Perfect score: 3792

Sequence: 1 atggcagatgaatcgagaa.....aatctttagacgaacatga 3792

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 139433 seqs, 91850399 residues

Word size : 0

Total number of hits satisfying chosen parameters: 278866

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_NA\_New.\*

- 1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	21	0.6	251	5	US-09-540-210B-23868
c 2	21	0.6	271	5	US-09-540-210B-31943
c 3	20	0.5	15769	6	US-10-105-299-13179
c 4	19	0.5	262	5	US-09-540-210B-9536
c 5	18	0.5	1487	6	US-10-115-123-109
c 6	18	0.5	1516	6	US-10-103-299-11641
c 7	18	0.5	1525	6	US-10-115-123-110
c 8	18	0.5	1949	6	US-10-115-123-26
c 9	18	0.5	2495	6	US-10-105-299-253
c 10	18	0.5	3291	6	US-10-105-299-6763
c 11	18	0.5	3291	6	US-10-105-299-6764
c 12	18	0.5	3291	6	US-10-105-299-6768
c 13	18	0.5	3291	6	US-10-105-299-6769
c 14	18	0.5	32121	6	US-10-103-313-604
c 15	17	0.4	242	5	US-09-540-210B-12852
c 16	17	0.4	290	5	US-09-540-210B-34467
c 17	17	0.4	296	5	US-09-540-210B-64
c 18	17	0.4	307	5	US-09-789-189-1791
c 19	17	0.4	308	5	US-09-540-210B-34304
c 20	17	0.4	787	6	US-10-102-806-91
c 21	17	0.4	912	5	US-09-540-209B-903
c 22	17	0.4	1269	5	US-09-540-209B-346
c 23	17	0.4	1308	6	US-10-011-588-14
c 24	17	0.4	1940	6	US-10-105-299-7324
c 25	17	0.4	1948	6	US-10-105-299-501
c 26	17	0.4	3069	5	US-09-941-492-105

#### ALIGNMENTS

RESULT 1  
US-09-540-210B-23868/c  
; Sequence 23868, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Sellhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997  
; PRIOR APPLICATION NUMBER: 60/014,010  
; PRIOR FILING DATE: March 25, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/903,555  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/862,178  
; PRIOR FILING DATE: May 22, 1997  
; PRIOR APPLICATION NUMBER: 60/018,217  
; PRIOR FILING DATE: May 23, 1996  
; PRIOR APPLICATION NUMBER: 08/881,589  
; PRIOR FILING DATE: June 24, 1997  
; PRIOR APPLICATION NUMBER: 60/021,275  
; PRIOR FILING DATE: June 25, 1996  
; PRIOR APPLICATION NUMBER: 08/903,802  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/905,881  
; PRIOR FILING DATE: August 1, 1997  
; PRIOR APPLICATION NUMBER: 60/025,204

Sequence 20, Appl  
Sequence 46, Appl  
Sequence 15108, A  
Sequence 11721, A  
Sequence 9857, Ap  
Sequence 11723, A  
Sequence 9140, Ap  
Sequence 1359, Ap  
Sequence 23459, A  
Sequence 11968, A  
Sequence 7777, Ap  
Sequence 4093, Ap  
Sequence 21907, A  
Sequence 18520, A  
Sequence 9378, Ap  
Sequence 24067, A  
Sequence 31167, A  
Sequence 1002, Ap  
Sequence 5788, Ap

;; PRIOR FILING DATE: August 1, 1996  
;; PRIOR APPLICATION NUMBER: 08/903,471  
;; PRIOR FILING DATE: July 30, 1997  
;; PRIOR APPLICATION NUMBER: 60/025,478  
;; PRIOR FILING DATE: July 31, 1996  
;; PRIOR APPLICATION NUMBER: 08/903,556  
;; PRIOR FILING DATE: July 31, 1997  
;; PRIOR APPLICATION NUMBER: 60/025,217  
;; PRIOR FILING DATE: August 22, 1996  
;; PRIOR APPLICATION NUMBER: 08/937,142  
;; PRIOR FILING DATE: September 23, 1997  
;; PRIOR APPLICATION NUMBER: 60/026,598  
;; PRIOR FILING DATE: September 24, 1996  
;; PRIOR APPLICATION NUMBER: 08/960,746  
;; PRIOR FILING DATE: October 29, 1997  
;; PRIOR APPLICATION NUMBER: 60/030,144  
;; PRIOR FILING DATE: October 30, 1996  
;; PRIOR APPLICATION NUMBER: 08/826,847  
;; PRIOR FILING DATE: April 10, 1997  
;; PRIOR APPLICATION NUMBER: 60/015,533  
;; PRIOR FILING DATE: April 10, 1996  
;; PRIOR APPLICATION NUMBER: 08/755,524  
;; PRIOR FILING DATE: November 22, 1996  
;; PRIOR APPLICATION NUMBER: 60/007,495  
;; PRIOR FILING DATE: November 22, 1995  
;; PRIOR APPLICATION NUMBER: 09/021,031  
;; PRIOR FILING DATE: February 10, 1998  
;; PRIOR APPLICATION NUMBER: 60/039,325  
;; PRIOR FILING DATE: February 13, 1997  
;; PRIOR APPLICATION NUMBER: 09/035,172  
;; PRIOR FILING DATE: March 4, 1998  
;; PRIOR APPLICATION NUMBER: 60/040,431  
;; PRIOR FILING DATE: March 5, 1997  
;; PRIOR APPLICATION NUMBER: 09/041,894  
;; PRIOR FILING DATE: March 12, 1998  
;; PRIOR APPLICATION NUMBER: 60/040,199  
;; PRIOR FILING DATE: March 14, 1997  
;; PRIOR APPLICATION NUMBER: 09/050,817  
;; PRIOR FILING DATE: March 30, 1998  
;; PRIOR APPLICATION NUMBER: 60/043,792  
;; PRIOR FILING DATE: April 11, 1997  
;; PRIOR APPLICATION NUMBER: 09/074,999  
;; PRIOR FILING DATE: May 8, 1998  
;; PRIOR APPLICATION NUMBER: 60/048,431  
;; PRIOR FILING DATE: May 29, 1997  
;; PRIOR APPLICATION NUMBER: 09/107,592  
;; PRIOR FILING DATE: June 30, 1998  
;; PRIOR APPLICATION NUMBER: 60/052,751  
;; PRIOR FILING DATE: July 1, 1997  
;; PRIOR APPLICATION NUMBER: 09/094,079  
;; PRIOR FILING DATE: June 9, 1998  
;; PRIOR APPLICATION NUMBER: 60/049,975  
;; PRIOR FILING DATE: June 13, 1997  
;; NUMBER OF SEQ ID NOS: 35654  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 23868  
;; LENGTH: 251  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: Incyte ID No: hu00855921  
US-09-540-210B-23868

Query Match 0.6%; Score 21; DB 5; Length 251;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1134 ccaaaacctggcatttgcta 1154  
Db 36 CCAAAACCTGGCATTTCCTA 16

RESULT 2  
US-09-540-210B-31943/c  
; Sequence 31943, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997  
; PRIOR APPLICATION NUMBER: 60/014,010  
; PRIOR FILING DATE: March 25, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/903,555  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/862,178  
; PRIOR FILING DATE: May 22, 1997  
; PRIOR APPLICATION NUMBER: 60/018,217  
; PRIOR FILING DATE: May 23, 1996  
; PRIOR APPLICATION NUMBER: 08/881,589  
; PRIOR FILING DATE: June 24, 1997  
; PRIOR APPLICATION NUMBER: 60/021,275  
; PRIOR FILING DATE: June 25, 1996  
; PRIOR APPLICATION NUMBER: 08/903,802  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/905,881  
; PRIOR FILING DATE: August 1, 1997  
; PRIOR APPLICATION NUMBER: 60/025,204  
; PRIOR FILING DATE: August 1, 1996  
; PRIOR APPLICATION NUMBER: 08/903,471  
; PRIOR FILING DATE: July 30, 1997  
; PRIOR APPLICATION NUMBER: 60/025,478  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/903,556  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/025,217  
; PRIOR FILING DATE: August 22, 1996  
; PRIOR APPLICATION NUMBER: 08/937,142  
; PRIOR FILING DATE: September 23, 1997  
; PRIOR APPLICATION NUMBER: 60/026,598  
; PRIOR FILING DATE: September 24, 1996  
; PRIOR APPLICATION NUMBER: 08/960,746  
; PRIOR FILING DATE: October 29, 1997  
; PRIOR APPLICATION NUMBER: 60/030,144  
; PRIOR FILING DATE: October 30, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/755,524  
; PRIOR FILING DATE: November 22, 1996



;> PRIOR APPLICATION NUMBER: 60/007,495  
;> PRIOR FILING DATE: November 22, 1995  
;> PRIOR APPLICATION NUMBER: 09/021,031  
;> PRIOR FILING DATE: February 10, 1998  
;> PRIOR APPLICATION NUMBER: 60/039,325  
;> PRIOR FILING DATE: February 13, 1997  
;> PRIOR APPLICATION NUMBER: 09/035,172  
;> PRIOR FILING DATE: March 4, 1998  
;> PRIOR APPLICATION NUMBER: 60/040,431  
;> PRIOR FILING DATE: March 5, 1997  
;> PRIOR APPLICATION NUMBER: 09/041,894  
;> PRIOR FILING DATE: March 12, 1998  
;> PRIOR APPLICATION NUMBER: 60/040,199  
;> PRIOR FILING DATE: March 14, 1997  
;> PRIOR APPLICATION NUMBER: 09/050,817  
;> PRIOR FILING DATE: March 30, 1998  
;> PRIOR APPLICATION NUMBER: 60/043,792  
;> PRIOR FILING DATE: April 11, 1997  
;> PRIOR APPLICATION NUMBER: 09/074,999  
;> PRIOR FILING DATE: May 8, 1998  
;> PRIOR APPLICATION NUMBER: 60/048,431  
;> PRIOR FILING DATE: May 29, 1997  
;> PRIOR APPLICATION NUMBER: 09/107,592  
;> PRIOR FILING DATE: June 30, 1998  
;> PRIOR APPLICATION NUMBER: 60/052,751  
;> PRIOR FILING DATE: July 1, 1997  
;> PRIOR APPLICATION NUMBER: 09/094,079  
;> PRIOR FILING DATE: June 9, 1998  
;> PRIOR APPLICATION NUMBER: 60/049,975  
;> PRIOR FILING DATE: June 13, 1997  
;> NUMBER OF SEQ ID NOS: 35654  
;> SOFTWARE: PERL Program  
;> SEQ ID NO 31943  
;> LENGTH: 271  
;> TYPE: DNA  
;> ORGANISM: Homo sapiens  
;> FEATURE:  
;> NAME/KEY: misc-feature  
;> OTHER INFORMATION: Incyte ID No: hu00652829  
US-09-540-210B-31943

Query Match 0.6%; Score 21; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1134 ccaaaacctggcattgcta 1154  
|||||  
Db 36 CCAAAACCTGGCATTGGCTA 16

RESULT 3  
US-10-105-299-13179/c  
;> Sequence 13179, Application us/10105299  
;> GENERAL INFORMATION:  
;> APPLICANT: Rosen, et. al  
;> TITLE OF INVENTION: Human Secreted Proteins  
;> FILE REFERENCE: PS950  
;> CURRENT APPLICATION NUMBER: US/10/105,299  
;> CURRENT FILING DATE: 2002-03-26  
;> NUMBER OF SEQ ID NOS: 15197  
;> Prior Application removed - See File Wrapper or Palm  
;> SOFTWARE: PatentIn Ver. 2.0  
;> SEQ ID NO 13179  
;> LENGTH: 15769  
;> TYPE: DNA  
;> ORGANISM: Homo sapiens  
US-10-105-299-13179

Query Match 0.5%; Score 20; DB 6; Length 15769;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1951 gaagcgaactcacacactac 1970  
Db 929 GAGCGCACTCACACCACTGC 910  
RESULT 4  
US-09-540-210B-9536/G  
;> Sequence 9536, Application US/09540210B  
;> GENERAL INFORMATION:  
;> APPLICANT: Seilhamer, Jeffrey J.  
;> APPLICANT: Deleageane, Angelo M.  
;> APPLICANT: Stuart, Susan G.  
;> APPLICANT: Stuve, Laura L.  
;> APPLICANT: Mullahy, Sara J.  
;> APPLICANT: Naughton, Rebecca E.  
;> TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
;> FILE REFERENCE: PD-1037 CIP  
;> CURRENT APPLICATION NUMBER: US/09/540,210B  
;> CURRENT FILING DATE: 2002-04-03  
;> PRIOR APPLICATION NUMBER: 08/972,899  
;> PRIOR FILING DATE: November 18, 1997  
;> PRIOR APPLICATION NUMBER: 08/395,244  
;> PRIOR FILING DATE: February 27, 1995  
;> PRIOR APPLICATION NUMBER: 08/722,922  
;> PRIOR FILING DATE: September 27, 1996  
;> PRIOR APPLICATION NUMBER: 60/005,526  
;> PRIOR FILING DATE: September 29, 1995  
;> PRIOR APPLICATION NUMBER: 08/824,029  
;> PRIOR FILING DATE: March 25, 1997  
;> PRIOR APPLICATION NUMBER: 60/014,010  
;> PRIOR FILING DATE: March 25, 1996  
;> PRIOR APPLICATION NUMBER: 08/826,847  
;> PRIOR FILING DATE: April 10, 1997  
;> PRIOR APPLICATION NUMBER: 60/015,533  
;> PRIOR FILING DATE: April 10, 1996  
;> PRIOR APPLICATION NUMBER: 08/903,555  
;> PRIOR FILING DATE: July 31, 1997  
;> PRIOR APPLICATION NUMBER: 60/023,308  
;> PRIOR FILING DATE: July 31, 1996  
;> PRIOR APPLICATION NUMBER: 08/862,178  
;> PRIOR FILING DATE: May 22, 1997  
;> PRIOR APPLICATION NUMBER: 60/018,217  
;> PRIOR FILING DATE: May 23, 1996  
;> PRIOR APPLICATION NUMBER: 08/881,589  
;> PRIOR FILING DATE: June 24, 1997  
;> PRIOR APPLICATION NUMBER: 60/021,275  
;> PRIOR FILING DATE: June 25, 1996  
;> PRIOR APPLICATION NUMBER: 08/903,802  
;> PRIOR FILING DATE: July 31, 1997  
;> PRIOR APPLICATION NUMBER: 60/023,308  
;> PRIOR FILING DATE: July 31, 1996  
;> PRIOR APPLICATION NUMBER: 08/905,881  
;> PRIOR FILING DATE: August 1, 1997  
;> PRIOR APPLICATION NUMBER: 60/025,204  
;> PRIOR FILING DATE: August 1, 1996  
;> PRIOR APPLICATION NUMBER: 08/903,471  
;> PRIOR FILING DATE: July 30, 1997  
;> PRIOR APPLICATION NUMBER: 60/025,478  
;> PRIOR FILING DATE: July 31, 1996  
;> PRIOR APPLICATION NUMBER: 08/903,556  
;> PRIOR FILING DATE: July 31, 1997  
;> PRIOR APPLICATION NUMBER: 60/025,217  
;> PRIOR FILING DATE: August 22, 1996  
;> PRIOR APPLICATION NUMBER: 08/937,142  
;> PRIOR FILING DATE: September 23, 1997  
;> PRIOR APPLICATION NUMBER: 60/026,598  
;> PRIOR FILING DATE: September 24, 1996  
;> PRIOR APPLICATION NUMBER: 08/960,746  
;> PRIOR FILING DATE: October 29, 1997  
;> PRIOR APPLICATION NUMBER: 60/030,144  
;> PRIOR FILING DATE: October 30, 1996  
;> PRIOR APPLICATION NUMBER: 08/826,847

; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/755,524  
; PRIOR FILING DATE: November 22, 1996  
; PRIOR APPLICATION NUMBER: 60/007,495  
; PRIOR FILING DATE: November 22, 1995  
; PRIOR APPLICATION NUMBER: 09/021,031  
; PRIOR FILING DATE: February 10, 1998  
; PRIOR APPLICATION NUMBER: 60/039,325  
; PRIOR FILING DATE: February 13, 1997  
; PRIOR APPLICATION NUMBER: 09/035,172  
; PRIOR FILING DATE: March 4, 1998  
; PRIOR APPLICATION NUMBER: 60/040,431  
; PRIOR FILING DATE: March 5, 1997  
; PRIOR APPLICATION NUMBER: 09/041,894  
; PRIOR FILING DATE: March 12, 1998  
; PRIOR APPLICATION NUMBER: 60/040,199  
; PRIOR FILING DATE: March 14, 1997  
; PRIOR APPLICATION NUMBER: 09/050,817  
; PRIOR FILING DATE: March 30, 1998  
; PRIOR APPLICATION NUMBER: 60/043,792  
; PRIOR FILING DATE: April 11, 1997  
; PRIOR APPLICATION NUMBER: 09/074,999  
; PRIOR FILING DATE: May 8, 1998  
; PRIOR APPLICATION NUMBER: 60/048,431  
; PRIOR FILING DATE: May 29, 1997  
; PRIOR APPLICATION NUMBER: 09/107,592  
; PRIOR FILING DATE: June 30, 1998  
; PRIOR APPLICATION NUMBER: 60/052,751  
; PRIOR FILING DATE: July 1, 1997  
; PRIOR APPLICATION NUMBER: 09/094,079  
; PRIOR FILING DATE: June 9, 1998  
; PRIOR APPLICATION NUMBER: 60/049,975  
; PRIOR FILING DATE: June 13, 1997  
; NUMBER OF SEQ ID NOS: 35654  
; SOFTWARE: PERL Program  
; SEQ ID NO 9536  
; LENGTH: 262  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: hu00667909  
US-09-540-210B-9536

Query Match 0.5%; Score 19; DB 5; Length 262;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1195 ttgacaattccagctggca 1213  
|||||  
Db 204 TTGACAATTCAGCTGGCA 186

RESULT 5  
US-10-115-123-109/c  
; Sequence 109, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16

; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 109  
; LENGTH: 1487  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (78)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (948)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-115-123-109

Query Match 0.5%; Score 18; DB 6; Length 1487;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 aggccaggatcaccagga 685  
|||||  
Db 570 AGGCCAGGATCACCAGGA 553

RESULT 6  
US-10-105-299-11641/c  
; Sequence 11641, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; PRIOR APPLICATION removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11641  
; LENGTH: 1516  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-11641

Query Match 0.5%; Score 18; DB 6; Length 1516;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 ttcaacttgacaattcca 1206  
|||||  
Db 1019 TTCAACTTGACAATTCCA 1002

RESULT 7  
US-10-115-123-110/c  
; Sequence 110, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508

; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 110  
; LENGTH: 1525  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (78)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-115-123-110

Query Match 0.5%; Score 18; DB 6; Length 1525;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 aggccaggatcacacagga 685  
|||||  
DB 570 AGCCAGGATCACCAGGA 553

RESULT 8  
US-10-115-123-26/c  
; Sequence 26, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 1949  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1130)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1948)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-115-123-26

Query Match 0.5%; Score 18; DB 6; Length 1949;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 aggccaggatcacacagga 685  
|||||  
DB 986 AGCCAGGATCACCAGGA 969

RESULT 9  
US-10-105-299-253  
; Sequence 253, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 253  
; LENGTH: 2495  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (913)..(913)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-105-299-253

Query Match 0.5%; Score 18; DB 6; Length 2495;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1199 caattccagctgggaaga 1216  
|||||  
DB 729 caattccagctgggaaga 746

RESULT 10  
US-10-105-299-6763  
; Sequence 6763, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6763  
; LENGTH: 3291  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-6763

Query Match 0.5%; Score 18; DB 6; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1199 caattccagctgggaaga 1216  
|||||  
DB 1557 caattccagctgggaaga 1574

RESULT 11  
US-10-105-299-6764  
; Sequence 6764, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950

; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6764  
; LENGTH: 3291  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-6764

Query Match 0.5%; Score 18; DB 6; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1199 caattccagctggcaaga 1216  
|||||  
Db 1557 caattccagctggcaaga 1574

RESULT 12  
US-10-105-299-6768  
; Sequence 6768, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6768  
; LENGTH: 3291  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-6768

Query Match 0.5%; Score 18; DB 6; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1199 caattccagctggcaaga 1216  
|||||  
Db 1557 caattccagctggcaaga 1574

RESULT 13  
US-10-105-299-6769  
; Sequence 6769, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6769  
; LENGTH: 3291  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-6769

Query Match 0.5%; Score 18; DB 6; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1199 caattccagctggcaaga 1216

Db 1557 caattccagctggcaaga 1574  
|||||

RESULT 14  
US-10-103-313-604/c  
; Sequence 604, Application US/10103313  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PJZ07C1  
; CURRENT APPLICATION NUMBER: US/10/103,313  
; CURRENT FILING DATE: 2002-03-12  
; NUMBER OF SEQ ID NOS: 653  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 604  
; LENGTH: 32121  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-103-313-604

Query Match 0.5%; Score 18; DB 6; Length 32121;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3534 aaagctgttgcacttga 3551  
|||||  
Db 30811 AAAGCTGTGCTACTTGA 30794

RESULT 15  
US-09-540-2108-12852/G  
; Sequence 12852, Application US/095402108  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullishy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,2108  
; CURRENT FILING DATE: 2002-04-03  
; Prior Application Number: 08/972,899  
; Prior Filing Date: November 18, 1997  
; Prior Application Number: 08/395,244  
; Prior Filing Date: February 27, 1995  
; Prior Application Number: 08/722,922  
; Prior Filing Date: September 27, 1996  
; Prior Application Number: 60/005,526  
; Prior Filing Date: September 23, 1995  
; Prior Application Number: 08/824,029  
; Prior Filing Date: March 25, 1997  
; Prior Application Number: 60/014,010  
; Prior Filing Date: March 25, 1996  
; Prior Application Number: 08/826,847  
; Prior Filing Date: April 10, 1997  
; Prior Application Number: 60/015,533  
; Prior Filing Date: April 10, 1996  
; Prior Application Number: 08/903,555  
; Prior Filing Date: July 31, 1997  
; Prior Application Number: 60/023,308  
; Prior Filing Date: July 31, 1996  
; Prior Application Number: 08/862,178  
; Prior Filing Date: May 22, 1997  
; Prior Application Number: 60/018,217  
; Prior Filing Date: May 23, 1996  
; Prior Application Number: 08/881,589  
; Prior Filing Date: June 24, 1997  
; Prior Application Number: 60/021,275

1 PRIOR FILING DATE: June 25, 1996  
2 PRIOR APPLICATION NUMBER: 08/903,802  
3 PRIOR FILING DATE: July 31, 1997  
4 PRIOR APPLICATION NUMBER: 60/023,308  
5 PRIOR FILING DATE: July 31, 1996  
6 PRIOR APPLICATION NUMBER: 08/905,881  
7 PRIOR FILING DATE: August 1, 1997  
8 PRIOR APPLICATION NUMBER: 60/025,204  
9 PRIOR FILING DATE: August 1, 1996  
10 PRIOR APPLICATION NUMBER: 08/903,471  
11 PRIOR FILING DATE: July 30, 1997  
12 PRIOR APPLICATION NUMBER: 60/025,478  
13 PRIOR FILING DATE: July 31, 1996  
14 PRIOR APPLICATION NUMBER: 08/903,556  
15 PRIOR FILING DATE: July 31, 1997  
16 PRIOR APPLICATION NUMBER: 60/025,217  
17 PRIOR FILING DATE: August 22, 1996  
18 PRIOR APPLICATION NUMBER: 08/937,142  
19 PRIOR FILING DATE: September 23, 1997  
20 PRIOR APPLICATION NUMBER: 60/026,598  
21 PRIOR FILING DATE: September 24, 1996  
22 PRIOR APPLICATION NUMBER: 08/960,746  
23 PRIOR FILING DATE: October 29, 1997  
24 PRIOR APPLICATION NUMBER: 60/030,144  
25 PRIOR FILING DATE: October 30, 1996  
26 PRIOR APPLICATION NUMBER: 08/826,847  
27 PRIOR FILING DATE: April 10, 1997  
28 PRIOR APPLICATION NUMBER: 60/015,533  
29 PRIOR FILING DATE: April 10, 1996  
30 PRIOR APPLICATION NUMBER: 08/755,524  
31 PRIOR FILING DATE: November 22, 1996  
32 PRIOR APPLICATION NUMBER: 60/007,495  
33 PRIOR FILING DATE: November 22, 1995  
34 PRIOR APPLICATION NUMBER: 09/021,031  
35 PRIOR FILING DATE: February 10, 1998  
36 PRIOR APPLICATION NUMBER: 60/039,325  
37 PRIOR FILING DATE: February 13, 1997  
38 PRIOR APPLICATION NUMBER: 09/035,172  
39 PRIOR FILING DATE: March 4, 1998  
40 PRIOR APPLICATION NUMBER: 60/040,431  
41 PRIOR FILING DATE: March 5, 1997  
42 PRIOR APPLICATION NUMBER: 09/041,894  
43 PRIOR FILING DATE: March 12, 1998  
44 PRIOR APPLICATION NUMBER: 60/040,199  
45 PRIOR FILING DATE: March 14, 1997  
46 PRIOR APPLICATION NUMBER: 09/050,817  
47 PRIOR FILING DATE: March 30, 1998  
48 PRIOR APPLICATION NUMBER: 60/043,792  
49 PRIOR FILING DATE: April 11, 1997  
50 PRIOR APPLICATION NUMBER: 09/074,999  
51 PRIOR FILING DATE: May 8, 1998  
52 PRIOR APPLICATION NUMBER: 60/048,431  
53 PRIOR FILING DATE: May 29, 1997  
54 PRIOR APPLICATION NUMBER: 09/107,592  
55 PRIOR FILING DATE: June 30, 1998  
56 PRIOR APPLICATION NUMBER: 60/052,751  
57 PRIOR FILING DATE: July 1, 1997  
58 PRIOR APPLICATION NUMBER: 09/094,079  
59 PRIOR FILING DATE: June 9, 1998  
60 PRIOR APPLICATION NUMBER: 60/049,975  
61 PRIOR FILING DATE: June 13, 1997  
62 NUMBER OF SEQ ID NOS: 35654  
63 SOFTWARE: PERL Program  
64 SEQ ID NO 12852  
65 LENGTH: 242  
66 TYPE: DNA  
67 ORGANISM: Homo sapiens  
68 FEATURE:  
69 NAME/KEY: misc\_feature  
70 OTHER INFORMATION: Incyte ID No: hu00604639  
71 US-09-540-210B-12852

Query Match 0.4%; Score 17; DB 5; Length 242;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1980 agctggagacctggaga 1996  
      |||||  
Db 184 ACCTGGAGACCTGGAGA 168

Search completed: April 27, 2002, 07:46:12  
Job time: 17224 sec

us-09-882-694a-10.rnpn

Mon Apr 29 09:00:17 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 07:42:58 ; Search time 8172.53 Seconds  
(without alignments)  
10038.450 Million cell updates/sec

Title: us-09-882-694a-10

Perfect score: 3792

Sequence: 1 atggcagatgaatcgagaa.....aatcttagaccagcatga 3792

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 21979536 seqs, 10817449327 residues

Word size : 0

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/pctus\_COMB.seq:\*

2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*

3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*

4: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq:\*

5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*

6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*

7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*

8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*

9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*

10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*

11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*

12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*

13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*

14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*

15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*

16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq:\*

17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*

18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*

19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*

20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*

21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*

22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*

23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:\*

24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:\*

25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:\*

26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:\*

27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:\*

28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*

29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*

30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*

31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:\*

32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq:\*

33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:\*

34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:\*

35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq:\*

36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:\*

37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq:\*

38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq:\*

39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq:\*

40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*

41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*

42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*

43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*

44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq:\*

45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq:\*

46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq:\*

47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq:\*

48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq:\*

49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq:\*

50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq:\*

51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq:\*

52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq:\*

53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq:\*

54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq:\*

55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:\*

56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq:\*

57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq:\*

58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq:\*

59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq:\*

60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:\*

61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq:\*

62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq:\*

63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq:\*

64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq:\*

65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq:\*

66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq:\*

67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq:\*

68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq:\*

69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq:\*

70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq:\*

71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq:\*

72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq:\*

73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq:\*

74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq:\*

75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3792	100.0	3792	17	US-09-351-224-10
2	3792	100.0	3792	17	US-09-351-224E-10
3	3792	100.0	3792	17	US-09-351-823-10
4	3792	100.0	3792	26	US-09-677-488-10
5	3792	100.0	3792	26	US-09-677-488A-10
6	3792	100.0	3792	26	US-09-677-682-10
7	3792	100.0	3792	26	US-09-677-682A-10
8	3792	100.0	3792	26	US-09-677-682B-10
9	3792	100.0	3792	33	US-09-882-694-10
10	3792	100.0	3792	33	US-09-882-694A-10
11	780	20.6	3999	17	US-09-351-224-9
12	780	20.6	3999	17	US-09-351-224E-9
13	780	20.6	3999	17	US-09-351-823-9
14	780	20.6	3999	26	US-09-677-488-9
15	780	20.6	3999	26	US-09-677-488A-9
16	780	20.6	3999	26	US-09-677-682-9
17	780	20.6	3999	26	US-09-677-682A-9
18	780	20.6	3999	26	US-09-677-682B-9
19	780	20.6	3999	33	US-09-882-694-9
20	780	20.6	3999	33	US-09-882-694A-9
21	22	0.6	848	19	US-09-505-532-4135
22	22	0.6	848	31	US-09-819-091A-4135
23	22	0.6	957	24	US-09-634-306B-179230
24	22	0.6	279270	61	US-60-226-176-556
25	22	0.6	279270	62	US-60-233-468-556
26	22	0.6	279270	70	US-60-313-371-556
27	21	0.6	251	13	US-08-937-142-2539
28	21	0.6	271	13	US-08-937-142-2519
29	21	0.6	275	15	US-09-129-873-1249
30	21	0.6	275	20	US-09-532-315-41139
31	21	0.6	335	52	US-60-132-861-1869





Qy 1561 ggaagcttgagtgaggtcaacacacgaattgcaatcgacagagtgatcgtatcgat 1620  
Db 1561 ggaagcttgagtgaggtcaacacacgaattgcaatcgacagagtgatcgtatcgat 1620  
Qy 1621 cccaagatcctgttactcgatgaagctaccagcgccctgaaccgagcgagaaagt 1680  
Db 1621 cccaagatcctgttactcgatgaagctaccagcgccctgaaccgagcgagaaagt 1680  
Qy 1681 gtccagagagccttgaaaccgagtgccaagacgcactacttggctgattgccacaaa 1740  
Db 1681 gtccagagagccttgaaaccgagtgccaagacgcactacttggctgattgccacaaa 1740  
Qy 1741 ctagccactgtccaaaagtgtgcaacatcgagtcatttcccaaggggaaaaatcgtag 1800  
Db 1741 ctagccactgtccaaaagtgtgcaacatcgagtcatttcccaaggggaaaaatcgtag 1800  
Qy 1801 caaggcacacaccgcaatgtgcaattcggtgtgctcattacgcccgcactggtgca 1860  
Db 1801 caaggcacacaccgcaatgtgcaattcggtgtgctcattacgcccgcactggtgca 1860  
Qy 1861 caggacctcggggtgacgaacacaaagacatgagaagaccctgcacgaagcgagca 1920  
Db 1861 caggacctcggggtgacgaacacaaagacatgagaagaccctgcacgaagcgagca 1920  
Qy 1921 cgagaagctgctgtgaacacggcaacttgagcgcaactcacaccactgccacatcaa 1980  
Db 1921 cgagaagctgctgtgaacacggcaacttgagcgcaactcacaccactgccacatcaa 1980  
Qy 1981 gctggagacctgggaagcggaaggtgcgggtcggaacttgggtactcgctctctaaaa 2040  
Db 1981 gctggagacctgggaagcggaaggtgcgggtcggaacttgggtactcgctctctaaaa 2040  
Qy 2041 tgcactcactcaatgttctacgaacaaaaatctctactggctctgttggctgaaca 2100  
Db 2041 tgcactcactcaatgttctacgaacaaaaatctctactggctctgttggctgaaca 2100  
Qy 2101 ataaaggtcttgatagtcggggccaaattccagggaagacccttgggttggagattg 2160  
Db 2101 ataaaggtcttgatagtcggggccaaattccagggaagacccttgggttggagattg 2160  
Qy 2161 ctcaactgtctcgagtgagtggtcatgcggaagacggcggaacttttatatttg 2220  
Db 2161 ctcaactgtctcgagtgagtggtcatgcggaagacggcggaacttttatatttg 2220  
Qy 2221 atgttcttggctgaggtctagaaatctagtagatttcaagatttgctggacatgc 2280  
Db 2221 atgttcttggctgaggtctagaaatctagtagatttcaagatttgctggacatgc 2280  
Qy 2281 aacgttattcacaagttgtacccatcgctatacaagcgcaatgttccaaagagtagtg 2340  
Db 2281 aacgttattcacaagttgtacccatcgctatacaagcgcaatgttccaaagagtagtg 2340  
Qy 2341 gatcaagacatcgaaactcctgcacatcccgagcaaatctggtctcacaatcgaa 2400  
Db 2341 gatcaagacatcgaaactcctgcacatcccgagcaaatctggtctcacaatcgaa 2400  
Qy 2401 ctgtcagctctaccacagcagttgcaagattgatatcagcaaatcttctcatttatc 2460  
Db 2401 ctgtcagctctaccacagcagttgcaagattgatatcagcaaatcttctcatttatc 2460  
Qy 2461 gtgtcggttcaacatcgctcgagcagtgctctaccactagcctatggatggaaactgg 2520  
Db 2461 gtgtcggttcaacatcgctcgagcagtgctctaccactagcctatggatggaaactgg 2520  
Qy 2521 cctgt 2580  
Db 2521 cctgt 2580  
Qy 2581 tctagagacgaagcttagaagcggaaactcggaacttggcagaagtgctgggtgc 2640  
Db 2581 tctagagacgaagcttagaagcggaaactcggaacttggcagaagtgctgggtgc 2640

Qy 2641 aagcgaagcagttaccgagatccggaccgtctcactcttggactctcgaagcgcatgtct 2700  
Db 2641 aagcgaagcagttaccgagatccggaccgtctcactcttggactctcgaagcgcatgtct 2700  
Qy 2701 ccaacagtaactcggacatgttgagcaaggtcttctaagatacacaagcttttgggtt 2760  
Db 2701 ccaacagtaactcggacatgttgagcaaggtcttctaagatacacaagcttttgggtt 2760  
Qy 2761 ggaagatgtttgttttctcacttgcacagtcgaaggagtttttggctatgacctggga 2820  
Db 2761 ggaagatgtttgttttctcacttgcacagtcgaaggagtttttggctatgacctggga 2820  
Qy 2821 tttgtattgcaatgataatgttcaagtgagtgagtaagacaacaactcaatttatata 2880  
Db 2821 tttgtattgcaatgataatgttcaagtgagtgagtaagacaacaactcaatttatata 2880  
Qy 2881 tcttcgtggcggttttgggttgcgggtccaaagcagcagccagttatttggcttactccag 2940  
Db 2881 tcttcgtggcggttttgggttgcgggtccaaagcagcagccagttatttggcttactccag 2940  
Qy 2941 agttttaccaggtcggctcggtcgaaactatatactctggtcggaacttgaagcg 3000  
Db 2941 agttttaccaggtcggctcggtcgaaactatatactctggtcggaacttgaagcg 3000  
Qy 3001 accatccggaacggagaggaacaaagggcccaagtcgggtggatgcctctgcac 3060  
Db 3001 accatccggaacggagaggaacaaagggcccaagtcgggtggatgcctctgcac 3060  
Qy 3061 ctgaggaacttaattcaggtatcgaactgaactggtcggtcggttctcggcggtt 3120  
Db 3061 ctgaggaacttaattcaggtatcgaactgaactggtcggtcggttctcggcggtt 3120  
Qy 3121 tccatgacaatcagcaggaacaaattttagcttattgtggcgctctcgtctggtgcaag 3180  
Db 3121 tccatgacaatcagcaggaacaaattttagcttattgtggcgctctcgtctggtgcaag 3180  
Qy 3181 tcaacgttgatcgttctgagcaacttctacgacccgacactcgggcggaatttcaatt 3240  
Db 3181 tcaacgttgatcgttctgagcaacttctacgacccgacactcgggcggaatttcaatt 3240  
Qy 3241 gcaacagagaatattgcagaaattgcgcgcgttctacccgagccatattgctttggtc 3300  
Db 3241 gcaacagagaatattgcagaaattgcgcgcgttctacccgagccatattgctttggtc 3300  
Qy 3301 caacaggaacccacactttaccaaggtccttcgcgagaaatgtgacgttggccctcgaa 3360  
Db 3301 caacaggaacccacactttaccaaggtccttcgcgagaaatgtgacgttggccctcgaa 3360  
Qy 3361 gccgaattatcagaagagcttttgaagagcgttctcgcgcgttctacccgagccatattgctttggtc 3420  
Db 3361 gccgaattatcagaagagcttttgaagagcgttctcgcgcgttctcgcgcgttcttggatt 3420  
Qy 3421 ttgtcactctttaccagaaggtttgaaacgcttgcggctcaacgagggatgcagttc 3480  
Db 3421 ttgtcactctttaccagaaggtttgaaacgcttgcggctcaacgagggatgcagttc 3480  
Qy 3481 tccggcggggaacgagcagatcgacatcgcaagcattgattcgaaatccaaagctg 3540  
Db 3481 tccggcggggaacgagcagatcgacatcgcaagcattgattcgaaatccaaagctg 3540  
Qy 3541 ttgtacttgacgaagcgagctcagccctgcacacgaatcgaacgctctggtttcaagct 3600  
Db 3541 ttgtacttgacgaagcgagctcagccctgcacacgaatcgaacgctctggtttcaagct 3600  
Qy 3601 gccctgataagcattccagcagccaaacgaatagcagtggtgcacagcatttccact 3660  
Db 3601 gccctgataagcattccagcagccaaacgaatagcagtggtgcacagcatttccact 3660  
Qy 3661 attcgggaatgtgattgttttttgggttgcgaacgggagaaatcgccgaacgggact 3720  
Db 3661 attcgggaatgtgattgttttttgggttgcgaacgggagaaatcgccgaacgggact 3720  
Qy 3721 caagcggaactacaacgacgagaggaagataattacgagatgtgtttgacaaatttta 3780

|||||  
Db 3721 cacgaggaactacacgactgagagagagataattacgagatgtgttggcacaatcttta 3780  
Qy 3781 gaccaagcatga 3792  
|||||  
Db 3781 gaccaagcatga 3792

RESULT 2  
US-09-351-224E-10  
: Sequence 10, Application US/09351224E  
: GENERAL INFORMATION:  
: APPLICANT: Duwick, Jon  
: APPLICANT: Maddox, Joyce  
: APPLICANT: Gilliam, Jacob  
: APPLICANT: Folkerts, Otto  
: APPLICANT: Crasta, Oswald R.  
: TITLE OF INVENTION: Compositions and Methods for Fumonisin  
: FILE OF INVENTION: Detoxification  
: CURRENT APPLICATION NUMBER: US/09/351.224E  
: CURRENT FILING DATE: 1999-07-12  
: NUMBER OF SEQ ID NOS: 11  
: SOFTWARE: FastSeq for Windows version 4.0  
: SEQ ID NO 10  
: LENGTH: 3792  
: TYPE: DNA  
: ORGANISM: Exophiala spinifera  
: FEATURE:  
: NAME/KEY: misc.feature  
: LOCATION: (0)...(0)  
: OTHER INFORMATION: P-glycoprotein, fully spliced cDNA  
US-09-351-224E-10

Query Match 100.0%; Score 3792; DB 17; Length 3792;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 atggcagatgaatcgagagaaactgcacaaacaaagatggcagtgatgctctcacac 60  
|||||  
Db 1 atggcagatgaatcgagagaaactgcacaaacaaagatggcagtgatgctctcacac 60  
Qy 61 cctcccccagaaaaagaaacgaagcagatattcagactatcagaaattcagatat 120  
|||||  
Db 61 cctcccccagaaaaagaaacgaagcagatattcagactatcagaaattcagatat 120  
Qy 121 gccgacaaatacagactggagactctcaatgtcatcgctcgtcatctcgccatcgatccggg 180  
|||||  
Db 121 gccgacaaatacagactggagactctcaatgtcatcgctcgtcatctcgccatcgatccggg 180  
Qy 181 gcttccttctctgatgtcgatcatcttcgtagcttcacaaacaaagttcaacaattac 240  
|||||  
Db 181 gcttccttctctgatgtcgatcatcttcgtagcttcacaaacaaagttcaacaattac 240  
Qy 241 aattcggcgacagggagtcctgaagcgtttcaaggccgagtggtggtatcttgcctgtgg 300  
Db 241 aattcggcgacagggagtcctgaagcgtttcaaggccgagtggtggtatcttgcctgtgg 300  
Qy 301 ttcgttaccttttattgggaagtgttctcactgactgtttccacggttcgcttacc 360  
|||||  
Db 301 ttcgttaccttttattgggaagtgttctcactgactgtttccacggttcgcttacc 360  
Qy 361 atttcagctatcagaaactcgaactcttcgagagtggttcttgaatgcacattcggg 420  
Db 361 atttcagctatcagaaactcgaactcttcgagagtggttcttgaatgcacattcggg 420  
Qy 421 caagaggtctggcatttcgacagcagagcaatgagcaatcgccactcargtcaactacc 480  
|||||  
Db 421 caagaggtctggcatttcgacagcagagcaatgagcaatcgccactcargtcaactacc 480  
Qy 481 aatggcaacccgtatatacaaacaggtatttcgagagaaattggtcttaccgtgcaggcaatt 540  
|||||

Db 481 aatggcaacccgtatatacaaacaggtatttcgagagaaattggtcttaccgtgcaggcaatt 540  
Qy 541 tcaatgtctcttcttgcatttgcgtcttgcgtcttgcgtcttgcgtcttgcgtcttgcgtcttgc 600  
Db 541 tcaatgtctcttcttgcatttgcgtcttgcgtcttgcgtcttgcgtcttgcgtcttgcgtcttgc 600  
Qy 601 accatgtccgtccatccgtccatttctcgtccgtccgtccgtccgtccgtccgtccgtccgtcc 660  
Db 601 accatgtccgtccatccgtccatttctcgtccgtccgtccgtccgtccgtccgtccgtccgtcc 660  
Qy 661 gctcaggagggcagagatcacccaggtatctacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720  
Db 661 gctcaggagggcagagatcacccaggtatctacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720  
Qy 721 ttatcatccatccagacagtcacgtcttctacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780  
Db 721 ttatcatccatccagacagtcacgtcttctacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780  
Qy 781 gatgtctttttgcagcaagcacacaaaggggagagaaatcgccaaataatggsctc 840  
Db 781 gatgtctttttgcagcaagcacacaaaggggagagaaatcgccaaataatggsctc 840  
Qy 841 ttgttctcaactgagatcttttgcatttgcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 900  
Db 841 ttgttctcaactgagatcttttgcatttgcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 900  
Qy 901 ttctgcattgcatcagaatggcaggttgc 960  
Db 901 ttctgcattgcatcagaatggcaggttgc 960  
Qy 961 tccgtcactcttagcagcagcagtcctcactcactcactcactcactcactcactcactcactcactc 1020  
Db 961 tccgtcactcttagcagcagcagtcctcactcactcactcactcactcactcactcactcactcactc 1020  
Qy 1021 caacgcccatcttcggtccggaattattcagtaataatgacaaacccagcagcagcagcagcagc 1080  
Db 1021 caacgcccatcttcggtccggaattattcagtaataatgacaaacccagcagcagcagcagcagc 1080  
Qy 1081 cctctcgacccttctggaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
Db 1081 cctctcgacccttctggaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
Qy 1141 ctggcatttgcctaccctccgcaccatctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1200  
Db 1141 ctggcatttgcctaccctccgcaccatctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1200  
Qy 1201 attccagctggcagagacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
Db 1201 attccagctggcagagacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
Qy 1261 gtcggttcttactgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Db 1261 gtcggttcttactgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Qy 1321 gaactgggacaaatacaatgtgaaatggtcagagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Db 1321 gaactgggacaaatacaatgtgaaatggtcagagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Qy 1381 cctgtgttcttctggtcagacaaatctccagacaaatctccagacaaatctccagacaaatctccagac 1440  
Db 1381 cctgtgttcttctggtcagacaaatctccagacaaatctccagacaaatctccagacaaatctccagac 1440  
Qy 1441 cagatctgcctcgc 1500  
Db 1441 cagatctgcctcgc 1500  
Qy 1501 ggcagcgttctcatttaataatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
Db 1501 ggcagcgttctcatttaataatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
Qy 1561 ggcagcgttctcatttaataatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620  
Db 1561 ggcagcgttctcatttaataatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620

QY 1621 cccaagatctctgttactcgatgaagtacacgagcccttgacccgaagcgcgagaaagtg 1680  
|||||  
Db 1621 cccaagatctctgttactcgatgaagtacacgagcccttgacccgaagcgcgagaaagtg 1680  
QY 1681 gtccaggagccctgaaccgaggtgccaaagaccgcaactacttggctaatgccacaag 1740  
|||||  
Db 1681 gtccaggagccctgaaccgaggtgccaaagaccgcaactacttggctaatgccacaag 1740  
QY 1741 ctgagccactgtcaaaagtgtcgcaacatcgagtcattccacggggaataatcgctcgag 1800  
|||||  
Db 1741 ctgagccactgtcaaaagtgtcgcaacatcgagtcattccacggggaataatcgctcgag 1800  
QY 1801 caaggcacacacccagaattgtatcgaaattcggtctgtcaatagccgcactgtgtcgatga 1860  
|||||  
Db 1801 caaggcacacacccagaattgtatcgaaattcggtctgtcaatagccgcactgtgtcgatga 1860  
QY 1861 caaggacctcggtgtgacgaacaaacaaacatgagaagaccctgcacgaagggcagca 1920  
|||||  
Db 1861 caaggacctcggtgtgacgaacaaacaaacatgagaagaccctgcacgaagggcagca 1920  
QY 1921 cgagaagctgtcgtgtgaacgacgcgcacttgagcgcactcacaccactgccacatctcaa 1980  
|||||  
Db 1921 cgagaagctgtcgtgtgaacgacgcgcacttgagcgcactcacaccactgccacatctcaa 1980  
QY 1981 gctgagacctggagaagcggaaggtgcggtgcgggactttgggtactcgtctcctaaaaa 2040  
|||||  
Db 1981 gctgagacctggagaagcggaaggtgcggtgcgggactttgggtactcgtctcctaaaaa 2040  
QY 2041 tgcattcctaatactattctacgaacaaaaaatctctactgctctctgttgtcaaca 2100  
|||||  
Db 2041 tgcattcctaatactattctacgaacaaaaaatctctactgctctctgttgtgtcaaca 2100  
QY 2101 ataaaggttctgatatgcgcggccacattctccaggagcaagccctttgttttcgagatg 2160  
|||||  
Db 2101 ataaaggttctgatatgcgcggccacattctccaggagcaagccctttgttttcgagatg 2160  
QY 2161 ctcaactgtcttcgagttgaagtgatgcggtgcggaagcaagcggtgagacttttatctcg 2220  
|||||  
Db 2161 ctcaactgtcttcgagttgaagtgatgcggtgcggaagcaagcggtgagacttttatctcg 2220  
QY 2221 atgtcttcttcgagttgaagttgtagatatttcaacgattggtggtgagatgc 2280  
|||||  
Db 2221 atgttcttcttcgagttgaagttgtagatatttcaacgattggtggtgagatgc 2280  
QY 2281 aacgttatttcacaagttgtccaccatcgtatcaagccgcaatgttccaaacgagtactg 2340  
|||||  
Db 2281 aacgttatttcacaagttgtccaccatcgtatcaagccgcaatgttccaaacgagtactg 2340  
QY 2341 gatcaagacatcgaaactctcgacatcccgaggaacaaatttctgtgctctcacatcgcaa 2400  
|||||  
Db 2341 gatcaagacatcgaaactctcgacatcccgaggaacaaatttctgtgctctcacatcgcaa 2400  
QY 2401 ctgtcagctctaccacgcagttgcaagagttgatatcagcaaaatttctcatttatc 2460  
|||||  
Db 2401 ctgtcagctctaccacgcagttgcaagagttgatatcagcaaaatttctcatttatc 2460  
QY 2461 gttgtcgtgtcaacatgctcagcagtgctctcacactagcctatggatggaaactggg 2520  
|||||  
Db 2461 gttgtcgtgtcaacatgctcagcagtgctctcacactagcctatggatggaaactggg 2520  
QY 2521 cctgt 2580  
|||||  
Db 2521 cctgt 2580  
QY 2581 tctagacgaagctagaaacggaactcgcaaaatttgcagaaggtctgtgtgtgtgtgtgtgt 2640  
|||||  
Db 2581 tctagacgaagctagaaacggaactcgcaaaatttgcagaaggtctgtgtgtgtgtgtgtgt 2640  
QY 2641 aagcggaagcagttacccgacatccgacgtctcacttttgactctcgaagccaatgtct 2700  
|||||  
Db 2641 aagcggaagcagttacccgacatccgacgtctcacttttgactctcgaagccaatgtct 2700

QY 2701 ccaacagttactcggacatgttgagcaaggtctctgtaagatcatccaagaattttggttt 2760  
|||||  
Db 2701 ccaacagttactcggacatgttgagcaaggtctctgtaagatcatccaagaattttggttt 2760  
QY 2761 ggacgatgttttggttttctactgtcagtcagtgaggtttttggctattgcctcgga 2820  
|||||  
Db 2761 ggacgatgttttggttttctactgtcagtcagtcagtgaggtttttggctattgcctcgga 2820  
QY 2821 tttgtattgagtcgataattggcttcaggtgagtcagacacaactcaatttatatca 2880  
|||||  
Db 2821 tttgtattgagtcgataattggcttcaggtgagtcagacacaactcaatttatatca 2880  
QY 2881 tttgtgtggcggtttttgttcgggtccaagcagcagcccagttattggcttactccacg 2940  
|||||  
Db 2881 tttgtgtggcggtttttgttcgggtccaagcagcagcccagttattggcttactccacg 2940  
QY 2941 agttttaccaaggtcgtcgtgcgaactatctctgtgctgcgagacattgaagcgc 3000  
|||||  
Db 2941 agttttaccaaggtcgtcgtgcgaactatctctgtgctgcgagacattgaagcgc 3000  
QY 3001 accatccgcgaacgggaggaacaaagaaaaagcccgatgggtggtgagtcgctgtcgac 3060  
|||||  
Db 3001 accatccgcgaacgggaggaacaaagaaaaagcccgatgggtggtgagtcgctgtcgac 3060  
QY 3061 ctcgaggacattgaattcagggtatcgtaacgtgatttcgggtcgagttctcgcggggtt 3120  
|||||  
Db 3061 ctcgaggacattgaattcagggtatcgtaacgtgatttcgggtcgagttctcgcggggtt 3120  
QY 3121 tccatgacaattcgagccagacaaatttgatcttatgtggcgcttctcgtgtggtgcaag 3180  
|||||  
Db 3121 tccatgacaattcgagccagacaaatttgatcttatgtggcgcttctcgtgtggtgcaag 3180  
QY 3181 tcaacgttgatcgtttgttcgggaacgattcttacgacccgacctcgggcccgaatttcatt 3240  
|||||  
Db 3181 tcaacgttgatcgtttgttcgggaacgattcttacgacccgacctcgggcccgaatttcatt 3240  
QY 3241 gcacagagaattatgcagaaatgtcgcgcgttggtacccgacctatgtctttgttc 3300  
|||||  
Db 3241 gcacagagaattatgcagaaatgtcgcgcgttggtacccgacctatgtctttgttc 3300  
QY 3301 caacaggaacccacactttaccaggtcgttcgcgagaaatgtgagcttggtccctcgaa 3360  
|||||  
Db 3301 caacaggaacccacactttaccaggtcgttcgcgagaaatgtgagcttggtccctcgaa 3360  
QY 3361 gccgaattatacgaagagctttgtcaaggacgccttcgcgagggccaatgctttggatt 3420  
|||||  
Db 3361 gccgaattatacgaagagctttgtcaaggacgccttcgcgagggccaatgctttggatt 3420  
QY 3421 ttgtcattctttaccagaagctttgaaacgccttcgcgctcaacgaggaatgcagttc 3480  
|||||  
Db 3421 ttgtcattctttaccagaagctttgaaacgccttcgcgctcaacgaggaatgcagttc 3480  
QY 3481 tccggtgggaacgacagcggatcgccatcgcaagagcattgattcgaaatccaaagctg 3540  
|||||  
Db 3481 tccggtgggaacgacagcggatcgccatcgcaagagcattgattcgaaatccaaagctg 3540  
QY 3541 ttgtacttgacgaagcgagtcagccctcgacacgcaatcggaacgtctggttcaagct 3600  
|||||  
Db 3541 ttgtacttgacgaagcgagtcagccctcgacacgcaatcggaacgtctggttcaagct 3600  
QY 3601 gccctcgatgagggatccagagccgaacgaataagcagtggtggcagcactttccact 3660  
|||||  
Db 3601 gccctcgatgagggatccagagccgaacgaataagcagtggtggcagcactttccact 3660  
QY 3661 attcggaatgtgatttattttgttttccaaacgggagaaatcgccgaacgggcaact 3720  
|||||  
Db 3661 attcggaatgtgatttattttgttttccaaacgggagaaatcgccgaacgggcaact 3720  
QY 3721 caacggaactacacgactgagaggaagattattacgagatgtgtttggcacaactttta 3780  
|||||  
Db 3721 caacggaactacacgactgagaggaagattattacgagatgtgtttggcacaactttta 3780  
QY 3781 gaccaagcagta 3792

```
Db 3781 gaccaagcatga 3792
|||||
RESULT 3
US-09-351-823-10
; Sequence 10, Application US/09351823
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification
; FILE REFERENCE: 5718-56
; CURRENT APPLICATION NUMBER: US/09/351,823
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3792
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; OTHER INFORMATION: p-glycoprotein, coding sequence
US-09-351-823-10

Query Match 100.0%; Score 3792; DB 17; Length 3792;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggcagatgaatcgagaaacctcgacaaacaaagatggcagtgagtcgctccacac 60
|||||
Db 1 atggcagatgaatcgagaaacctcgacaaacaaagatggcagtgagtcgctccacac 60
|||||
Qy 61 cctccccagaaaaagaaacggcagtgatttcagactactcagaaatcttcagatat 120
|||||
Db 61 cctccccagaaaaagaaacggcagtgatttcagactactcagaaatcttcagatat 120
|||||
Qy 121 gccgacaaatcacgactgactctcaatgtcgcgtcactcgcgcacatcgccatccggg 180
|||||
Db 121 gccgacaaatcacgactgactctcaatgtcgcgtcactcgcgcacatcgccatccggg 180
|||||
Qy 181 gcttcctctctgactgactcactcttcgtagcttcacaaacaaagttcaaaattac 240
|||||
Db 181 gcttcctctctgactgactcactcttcgtagcttcacaaacaaagttcaaaattac 240
|||||
Qy 241 aattcggcgacggagtcctcgaagcgttcaagggcgtatggatcatttcctgtgg 300
|||||
Db 241 aattcggcgacggagtcctcgaagcgttcaagggcgtatggatcatttcctgtgg 300
|||||
Qy 301 ttgcttacctctttattgggaagtgttgctcactcactgcttccacggctgccattacc 360
|||||
Db 301 ttgcttacctctttattgggaagtgttgctcactcactgcttccacggctgccattacc 360
|||||
Qy 361 atttcagatatcgaacactcgaactcttcgacgagtggttccttgaatgaacttgccg 420
|||||
Db 361 atttcagatatcgaacactcgaactcttcgacgagtggttccttgaatgaacttgccg 420
|||||
Qy 421 caagaggctctggcatttcgacaaagcagagcaatggagcaatcgccactcargtcaacc 480
|||||
Db 421 caagaggctctggcatttcgacaaagcagagcaatggagcaatcgccactcargtcaacc 480
|||||
Qy 481 aatggcaaccgtatatacaaacaggatttcccgagaaatttggctttaccgtgcaggcaatt 540
|||||
Db 481 aatggcaaccgtatatacaaacaggatttcccgagaaatttggctttaccgtgcaggcaatt 540
|||||
Qy 541 tcaatgtctctctctgacttctgttgctcttgcttcagtgagtgaaagctagctttaacc 600
|||||
Db 541 tcaatgtctctctctgacttctgttgctcttgcttcagtgagtgaaagctagctttaacc 600
|||||
Qy 601 accatgtccgtcaatccctgcacatttctcgttcacccgcatctgcatagcaattgatgcc 660
|||||
Db 601 accatgtccgtcaatccctgcacatttctcgttcacccgcatctgcatagcaattgatgcc 660
|||||
Qy 661 gctcaggaggccaggaatcaccaggaatctactcaacgctcgcgtgtcctcgcagaaagtc 720
|||||
```

```
Db 661 gctcaggaggccaggatcacccaggatctactcacgcccgtgtcctcgcagaaagtc 720
|||||
Qy 721 ttatcatcatccgacagtcacatgcttctacgcccagagaaaaatgggtcgaaaaatcat 780
|||||
Db 721 ttatcatcatccgacagtcacatgcttctacgcccagagaaaaatgggtcgaaaaatcat 780
|||||
Qy 781 gatgtctttttgacagcaagcacaccaagagggaagaaatcgccaaaaataatggsgtc 840
|||||
Db 781 gatgtctttttgacagcaagcacaccaagagggaagaaatcgccaaaaataatggsgtc 840
|||||
Qy 841 ttgttctcaactgagtgacttttgcatttaccgtgctactcgcactggccttttggaaaggt 900
|||||
Db 841 ttgttctcaactgagtgacttttgcatttaccgtgctactcgcactggccttttggaaaggt 900
|||||
Qy 901 ttctgcagtatcagaatggcgaggttgcgacgttggcaagtcttactgttgcctt 960
|||||
Db 901 ttctgcagtatcagaatggcgaggttgcgacgttggcaagtcttactgttgcctt 960
|||||
Qy 961 tccgtcaccttagcagccagtcacatcctcaatgcttgccttcaggttcagtcgtttac 1020
|||||
Db 961 tccgtcaccttagcagccagtcacatcctcaatgcttgccttcaggttcagtcgtttac 1020
|||||
Qy 1021 caacgcgcgaatcttcggctccgaattattcagtgatatattgacaaacccacgcagctcgac 1080
|||||
Db 1021 caacgcgcgaatcttcggctccgaattattcagtgatatattgacaaacccacgcagctcgac 1080
|||||
Qy 1081 cctctgcaccttctggaagcagccagagggctgcttaggtcaaaatggatccaaac 1140
|||||
Db 1081 cctctgcaccttctggaagcagccagagggctgcttaggtcaaaatggatccaaac 1140
|||||
Qy 1141 ctggcatttgcttaccctccgacacatctgccaaagtactctcagatttcaacttgaca 1200
|||||
Db 1141 ctggcatttgcttaccctccgacacatctgccaaagtactctcagatttcaacttgaca 1200
|||||
Qy 1201 attccagctggcgaagcagcgcctcgtcgttgctcaggtgcaatggcgaagcaaatg 1260
|||||
Db 1201 attccagctggcgaagcagcgcctcgtcgttgctcaggtgcaatggcgaagcaaatg 1260
|||||
Qy 1261 gtcggcttacttgaaacggttgatctgccagttcggggaggaattacttgatgggttg 1320
|||||
Db 1261 gtcggcttacttgaaacggttgatctgccagttcggggaggaattacttgatgggttg 1320
|||||
Qy 1321 gaactgggacaaacacaaatggaatgctgagaaagcgcattcgcctgttcaacagaaa 1380
|||||
Db 1321 gaactgggacaaacacaaatggaatgctgagaaagcgcattcgcctgttcaacagaaa 1380
|||||
Qy 1381 cctgtgtgttctgtggcacaatcttcacagaaatggcgaacggttcaatggatgagcaa 1440
|||||
Db 1381 cctgtgtgttctgtggcacaatcttcacagaaatggcgaacggttcaatggatgagcaa 1440
|||||
Qy 1441 cgagatctcctcgcgaaaaacaaatggagcttgcgcaaaagcttcgaagccagcaat 1500
|||||
Db 1441 cgagatctcctcgcgaaaaacaaatggagcttgcgcaaaagcttcgaagccagcaat 1500
|||||
Qy 1501 ggcgacgctgttcattaatgagcttcgaaacggttatagagactgaagtggcgagcgagcc 1560
|||||
Db 1501 ggcgacgctgttcattaatgagcttcgaaacggttatagagactgaagtggcgagcgagcc 1560
|||||
Qy 1561 gaagccttgagtgagggttcacacgacaaacgaatgcaatcgacgaaagtaataatcgagat 1620
|||||
Db 1561 gaagccttgagtgagggttcacacgacaaacgaatgcaatcgacgaaagtaataatcgagat 1620
|||||
Qy 1621 cccaagatcctgttactcgatgaagtaccagcgcccttgaccgagagcgaggaaggtg 1680
|||||
Db 1621 cccaagatcctgttactcgatgaagtaccagcgcccttgaccgagagcgaggaaggtg 1680
|||||
Qy 1681 gtccagagagccttgaaacggagtgccaaagacgcactacttttggctcattggccacaaa 1740
|||||
Db 1681 gtccagagagccttgaaacggagtgccaaagacgcactacttttggctcattggccacaaa 1740
|||||
Qy 1741 cttagccactgtcaaaagtgctggcacaacatcgacgtctatttccaggggaaaaatcgtcgag 1800
|||||
```

Db 1741 ctaggcactgtcaaaagtgctggcaacatcgagtcatttcccagggggaaaaatcgtcgag 1800  
QY 1801 caaggcacacacacgaattgacgaattcggctgtcattacgcgcgcactggctgcgtgca 1860  
Db 1801 caaggcacacacacgaattgacgaattcggctgtcattacgcgcgcactggctgcgtgca 1860  
QY 1861 caaggacctcgggctgacgaacaacaagaacatgagaagacctgcacgaaaaggcagca 1920  
Db 1861 cadgacctcgggctgacgaacaacaagaacatgagaagacctgcacgaaaaggcagca 1920  
QY 1921 cgagaagctcgtgtgaacacgcgcgcacttgagcgcactcacaccactgccacatctcaa 1980  
Db 1921 cgagaagctcgtgtgaacacgcgcgcacttgagcgcactcacaccactgccacatctcaa 1980  
QY 1981 gctggagacctgagaaagcgaaagtgccggtcgggactttgggctactcgtcctcaaaa 2040  
Db 1981 gctggagacctgagaaagcgaaagtgccggtcgggactttgggctactcgtcctcaaaa 2040  
QY 2041 tgcattcctaatactgttctacgaacaaaaaaatctctacttggtgctctcttctgttcaaca 2100  
Db 2041 tgcattcctaatactgttctacgaacaaaaaaatctctacttggtgctctcttctgttcaaca 2100  
QY 2101 ataacggttctgatatacgcgccacacatttccaggacaaagcccttttggtttctgagattg 2160  
Db 2101 ataacggttctgatatacgcgccacacatttccaggacaaagcccttttggtttctgagattg 2160  
QY 2161 ctcaactgtctcagttgagtggtcattgctgcgcgacagaaagcggcagacttttatattctg 2220  
Db 2161 ctcaactgtctcagttgagtggtcattgctgcgcgacagaaagcggcagacttttatattctg 2220  
QY 2221 atgttctttctggtcttaggaaatctagtaggatatatttcaagattgctgacatgc 2280  
Db 2221 atgttctttctggtcttaggaaatctagtaggatatatttcaagattgctgacatgc 2280  
QY 2281 aacgttatttcaaaagttgtcacccatcgctatcaagccgcaatgtttccaacgagtaactg 2340  
Db 2281 aacgttatttcaaaagttgtcacccatcgctatcaagccgcaatgtttccaacgagtaactg 2340  
QY 2341 gataaagacatcgaactcctcgcacatcccggagcaaatcttctgctgcctcacatcgcaa 2400  
Db 2341 gataaagacatcgaactcctcgcacatcccggagcaaatcttctgctgcctcacatcgcaa 2400  
QY 2401 ctgtcagctctaccacgcagttgcaagatttgatatcacaaaatttctcatttatctc 2460  
Db 2401 ctgtcagctctaccacgcagttgcaagatttgatatcacaaaatttctcatttatctc 2460  
QY 2461 gttgtcgttcaaacatcgtctcgagcagtgctctaccactagcctatgtaggaaactggg 2520  
Db 2461 gttgtcgttcaaacatcgtctcgagcagtgctctaccactagcctatgtaggaaactggg 2520  
QY 2521 cctggtggttgtgttgggtgcacttccaccctcgtcttgggtggttacctcagaattctg 2580  
Db 2521 cctggtggttgtgttgggtgcacttccaccctcgtcttgggtggttacctcagaattctg 2580  
QY 2581 tctagacgaagtagaagccggaactcggcgaactttgcagaagtgctggtcctgc 2640  
Db 2581 tctagacgaagtagaagccggaactcggcgaactttgcagaagtgctggtcctgc 2640  
QY 2641 aagcgaagcagttaccgcgactcggacgctctcattcttgcactctcgaaagccatgttct 2700  
Db 2641 aagcgaagcagttaccgcgactcggacgctctcattcttgcactctcgaaagccatgttct 2700  
QY 2701 ccaacagttactcggacatgtttgagcaaggtcttctaagatcatccaaagcttttgggtt 2760  
Db 2701 ccaacagttactcggacatgtttgagcaaggtcttctaagatcatccaaagcttttgggtt 2760  
QY 2761 ggaacagtttttgggttttccacttgcacagtcgatggagtttttggctattgcccctggga 2820  
Db 2761 ggaacagtttttgggttttccacttgcacagtcgatggagtttttggctattgcccctggga 2820  
QY 2821 ttgttattgcagtcgataattgcttcaggttcaggttcagtcgacacacactcaatttatatca 2880  
Db 2821 ttgttattgcagtcgataattgcttcaggttcaggttcagtcgacacacactcaatttatatca 2880

QY 2881 tcttcgtgggctgttctgttgcgggtccaaagcagcagcccagttatttggcttaactccag 2940  
Db 2881 tcttcgtgggctgttctgttgcgggtccaaagcagcagcccagttatttggcttaactccag 2940  
QY 2941 agttttaccaaggtcgcgtcggctcggaactatactctcgtcgcgcgacattgaagcgc 3000  
Db 2941 agttttaccaaggtcgcgtcggctcggaactatactctcgtcgcgcgacattgaagcgc 3000  
QY 3001 accatccgcgaacacgggaggaacaagaaaaagcccagtggtggatgcctgtcgac 3060  
Db 3001 accatccgcgaacacgggaggaacaagaaaaagcccagtggtggatgcctgtcgac 3060  
QY 3061 ctcgaggacattgaattcaggtatcgtcaacgttatcggtcgcgacttctccgcgggtt 3120  
Db 3061 ctcgaggacattgaattcaggtatcgtcaacgttatcggtcgcgacttctccgcgggtt 3120  
QY 3121 tccatgacaaatcgaccagggacaattttagcttatgtggcgcttctcgtcgtggaag 3180  
Db 3121 tccatgacaaatcgaccagggacaattttagcttatgtggcgcttctcgtcgtggaag 3180  
QY 3181 tcaacgttgatcgtttgtcgggaacgatttctacgaccgcacctcgggcccgaatttcattt 3240  
Db 3181 tcaacgttgatcgtttgtcgggaacgatttctacgaccgcacctcgggcccgaatttcattt 3240  
QY 3241 gcaacgagaataattgcagaaaatgtcgcgcgcgtctgtacccgcgcacatatgtcttggtc 3300  
Db 3241 gcaacgagaataattgcagaaaatgtcgcgcgcgtctgtacccgcgcacatatgtcttggtc 3300  
QY 3301 caacgaggaaccacacactttacaaagctcgttcgcgagaatgtgacgttggccctcgaa 3360  
Db 3301 caacgaggaaccacacactttacaaagctcgttcgcgagaatgtgacgttggccctcgaa 3360  
QY 3361 gccgaattatcagaagagctttgtcgaagagccttcccgcaaggccaatgctttggatt 3420  
Db 3361 gccgaattatcagaagagctttgtcgaagagccttcccgcaaggccaatgctttggatt 3420  
QY 3421 ttgtcatctcttaccagaagctttgaaaccccttgcgctcaacagaggtgcagttc 3480  
Db 3421 ttgtcatctcttaccagaagctttgaaaccccttgcgctcaacagaggtgcagttc 3480  
QY 3481 tccggcggaacacagcagcggatcgccatcgcaagagcattgattcgaatccaaagctg 3540  
Db 3481 tccggcggaacacagcagcggatcgccatcgcaagagcattgattcgaatccaaagctg 3540  
QY 3541 ttgtcatttgacgaagcagcgtcagccctcgacaacgaatcggaacgctggttcaagct 3600  
Db 3541 ttgtcatttgacgaagcagcgtcagccctcgacaacgaatcggaacgctggttcaagct 3600  
QY 3601 gccctcgatgagcattccacgagcgcgaacgaatagcagtggtgcccagacttccact 3660  
Db 3601 gccctcgatgagcattccacgagcgcgaacgaatagcagtggtgcccagacttccact 3660  
QY 3661 attcggaaatgtgatttatttctgttgcgaacggagaaatccggaacgggcact 3720  
Db 3661 attcggaaatgtgatttatttctgttgcgaacggagaaatccggaacgggcact 3720  
QY 3721 cagcgggaactacaacgactgagaggaatattacgagatgtgtttggcacaatttta 3780  
Db 3721 cagcgggaactacaacgactgagaggaatattacgagatgtgtttggcacaatttta 3780  
QY 3781 gaccaagcatga 3792  
Db 3781 gaccaagcatga 3792

RESULT 4  
US-09-677-488-10  
; Sequence 10, Application US/09677488  
; GENERAL INFORMATION:  
; APPLICANT: Duwick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 35718/204100 (5718-111A )

: CURRENT APPLICATION NUMBER: US/09/677,488									
: CURRENT FILING DATE: 2000-10-02									
: PRIOR APPLICATION NUMBER: 09/351,224									
: PRIOR FILING DATE: 1999-07-12									
: NUMBER OF SEQ ID NOS: 11									
: SOFTWARE: PatentIn Ver. 2.0									
: SEQ ID NO 10									
: LENGTH: 3792									
: TYPE: DNA									
: ORGANISM: Exophiala spinifera									
: FEATURE:									
: OTHER INFORMATION: p-glycoprotein, fully spliced cDNA									
US-09-677-488-10									
Query Match 100.0%; Score 3792; DB 26; Length 3792;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps									
Qy	1	atggcagatgaatcggagaacacctgcaccaacaagaatgcagtcgctcctcaac	60						
Db	1	atggcagatgaatcggagaacacctgcaccaacaagaatgcagtcgctcctcaac	60						
Qy	61	cctcccagaaaaggaaacccgaagcgattctcagactatctacgaattctcagatat	120						
Db	61	cctcccagaaaaggaaacccgaagcgattctcagactatctacgaattctcagatat	120						
Qy	121	gcgagaaaatcagactggactccaatgtcatcgctcatctcgcccatcgatccggg	180						
Db	121	gcgagaaaatcagactggactccaatgtcatcgctcatctcgcccatcgatccggg	180						
Qy	181	gcttcctctctctgattgcgatacattctcgttagcttcaccaacaagtccaacaattac	240						
Db	181	gcttcctctctctgattgcgatacattctcgttagcttcaccaacaagtccaacaattac	240						
Qy	241	aattcggcgacgggagtcctgaagcgttcaaggccgagtggatcatttgcctgtgtg	300						
Db	241	aattcggcgacgggagtcctgaagcgttcaaggccgagtggatcatttgcctgtgtg	300						
Qy	301	ttcgtctacactctttattgggaagttgttcctcacgtacgtttccacgctccattacc	360						
Db	301	ttcgtctacactctttattgggaagttgttcctcacgtacgtttccacgctccattacc	360						
Qy	361	attcagctatcagaacactcgaaactcttcgacgagttcttcgtgaatgcaccttg	420						
Db	361	attcagctatcagaacactcgaaactcttcgacgagttcttcgtgaatgcaccttg	420						
Qy	421	caagaggtctgcattctcacaagcagacaatcgagacaatcgcaactcarctactacc	480						
Db	421	caagaggtctgcattctcacaagcagacaatcgagacaatcgcaactcarctactacc	480						
Qy	481	aatggcaaccgtatacaaacaggtattgcgagaaattggtcttaccgtcaggcaact	540						
Db	481	aatggcaaccgtatacaaacaggtattgcgagaaattggtcttaccgtcaggcaact	540						
Qy	541	tcaatgtctcttcctcatttggcgtgtttggcgctcagtggaagtacgtttaatc	600						
Db	541	tcaatgtctcttcctcatttggcgtgtttggcgctcagtggaagtacgtttaatc	600						
Qy	601	accatgtccgtcatccctgcattcttccttggttcacggcatctcgatagaactgatgc	660						
Db	601	accatgtccgtcatccctgcattcttccttggttcacggcatctcgatagaactgatgc	660						
Qy	661	gctcaggagccaggatcacccaggaatctactcacgcgcgcgtctctcagagaagtc	720						
Db	661	gctcaggagccaggatcacccaggaatctactcacgcgcgcgtctctcagagaagtc	720						
Qy	721	tatcatccaatccgacagtccaatgcttctacgccagagaagaaatgctcgaaaaaat	780						
Db	721	tatcatccaatccgacagtccaatgcttctacgccagagaagaaatgctcgaaaaaat	780						
Qy	781	gatgtcttttttcgacgaacacacccaagaaggaggaagaataatcgccaataatgsgtc	840						
Db	781	gatgtcttttttcgacgaacacacccaagaaggaggaagaataatcgccaataatgsgtc	840						

QY 1921 cgaagactgctgggtgaacgacccgcaacttgagcgcaactcacacacactgccacatctcaa 1980  
DB 1921 cgaagactgctgggtgaacgacccgcaacttgagcgcaactcacacacactgccacatctcaa 1980  
QY 1981 gctggagacctggagaagcgaaggtgcgggtcgggactttgggctactcgcctctaaaa 2040  
DB 1981 gctggagacctggagaagcgaaggtgcgggtcgggactttgggctactcgcctctaaaa 2040  
QY 2041 tgcattcttaactatgtctctacgaaacaaaaaatctctactggtgcttctgttgcaaca 2100  
DB 2041 tgcattcttaactatgtctctacgaaacaaaaaatctctactggtgcttctgttgcaaca 2100  
QY 2101 ataacggttctgatatgcgcggccacatttccagacaaagcccttttggtcgagattg 2160  
DB 2101 ataacggttctgatatgcgcggccacatttccagacaaagcccttttggtcgagattg 2160  
QY 2161 ctcaactgtcttcgagttgagttggttcggtgcgcagaggaacggttttatattctg 2220  
DB 2161 ctcaactgtcttcgagttgagttggttcggtgcgcagaggaacggttttatattctg 2220  
QY 2221 atgttcttctgctggtctctaggaatactagtagatatatttcaacgattggctggacatgc 2280  
DB 2221 atgttcttctgctggtctctaggaatactagtagatatatttcaacgattggctggacatgc 2280  
QY 2281 aacgttatttcaacagttgtcaccatcgctatcaagcgcgaatgttccaaacgagtactg 2340  
DB 2281 aacgttatttcaacagttgtcaccatcgctatcaagcgcgaatgttccaaacgagtactg 2340  
QY 2341 gatcaagacatcgaaactcctcgacatcccggaagcaaatcttctggtctctacatcgcaa 2400  
DB 2341 gatcaagacatcgaaactcctcgacatcccggaagcaaatcttctggtctctacatcgcaa 2400  
QY 2401 ctgcagactctaccacgacgttgcaagattgatatcagcaaaatttctcatttatatc 2460  
DB 2401 ctgcagactctaccacgacgttgcaagattgatatcagcaaaatttctcatttatatc 2460  
QY 2461 gttctgggtcaacatcgtctctgagcagtgctctaccactagcctatgtagtgaactggg 2520  
DB 2461 gttctgggtcaacatcgtctctgagcagtgctctaccactagcctatgtagtgaactggg 2520  
QY 2521 cctgggtggtgtgttgggtgcaactccaccctgcttttgggtggtactccacgaattcg 2580  
DB 2521 cctgggtggtgtgttgggtgcaactccaccctgcttttgggtggtactccacgaattcg 2580  
QY 2581 tctagagacgagctagaagccgaaactcggcaactttgcagaaagtgcgtggcttgc 2640  
DB 2581 tctagagacgagctagaagccgaaactcggcaactttgcagaaagtgcgtggcttgc 2640  
QY 2641 aagcgaagcagttaccgcgattcccgaccgtctctactttgactctcgaaggccattgtct 2700  
DB 2641 aagcgaagcagttaccgcgattcccgaccgtctctactttgactctcgaaggccattgtct 2700  
QY 2701 ccaacagtactcgcgacatggttgagcaaggttctgtaagatcatccaaagcttttggttt 2760  
DB 2701 ccaacagtactcgcgacatggttgagcaaggttctgtaagatcatccaaagcttttggttt 2760  
QY 2761 ggacagtttttgggttttctacattcagctcagctgagtttttggctatgcccctggga 2820  
DB 2761 ggacagtttttgggttttctacattcagctcagctgagtttttggctatgcccctggga 2820  
QY 2821 ttttgtattgcagtcgaataattggtctcaggtgagtcgacacaaactcaattttatatca 2880  
DB 2821 ttttgtattgcagtcgaataattggtctcaggtgagtcgacacaaactcaattttatatca 2880  
QY 2881 tcttcgtggggttttgggttcgggttcgaagcagcagcccagttatttgggttaactccag 2940  
DB 2881 tcttcgtggggttttgggttcgggttcgaagcagcagcccagttatttgggttaactccag 2940  
QY 2941 agttttaccaggtctcggtcggtgcgaactatatctctggtcgcgacattgaagccg 3000  
DB 2941 agttttaccaggtctcggtcggtgcgaactatatctctggtcgcgacattgaagccg 3000

QY 3001 accatccgcgaaacgagcgaggagaaacaaaaagggccacagtgggtggtgacccctgtcgac 3060  
DB 3001 accatccgcgaaacgagcgaggagaaacaaaaagggccacagtgggtggtgacccctgtcgac 3060  
QY 3061 ctcgaggacaattgaattcaggtatcgtcaacgtgattcggctcgagttctctccgcggggtt 3120  
DB 3061 ctcgaggacaattgaattcaggtatcgtcaacgtgattcggctcgagttctctccgcggggtt 3120  
QY 3121 tccatgacaatcgagccagcgacaaattttagcttatgtgggcgtcttctggctggtgcaag 3180  
DB 3121 tccatgacaatcgagccagcgacaaattttagcttatgtgggcgtcttctggctggtgcaag 3180  
QY 3181 tcaacttgatcgctttgttcgggaacgattttagaccccgacctcgggccgaatttcaatt 3240  
DB 3181 tcaacttgatcgctttgttcgggaacgattttagaccccgacctcgggccgaatttcaatt 3240  
QY 3241 gcacacgagaataattgcagaaattgcgcgcgcttgcacgcggccaatatgtctttggtc 3300  
DB 3241 gcacacgagaataattgcagaaattgcgcgcgcttgcacgcggccaatatgtctttggtc 3300  
QY 3301 caacaggaacccacatttaccagaggtctcggttcggagaaatgtgacgttggccctcgaa 3360  
DB 3301 caacaggaacccacatttaccagaggtctcggttcggagaaatgtgacgttggccctcgaa 3360  
QY 3361 gcgaattatcagaagagctttgtcaaggacgccttcccgcaaggccaattgttggatt 3420  
DB 3361 gcgaattatcagaagagctttgtcaaggacgccttcccgcaaggccaattgttggatt 3420  
QY 3421 ttgtctatctctttaccagaaggtttgaaacgccttgcggtctcaacgagggatgcagttc 3480  
DB 3421 ttgtctatctctttaccagaaggtttgaaacgccttgcggtctcaacgagggatgcagttc 3480  
QY 3481 tccggcggaacacgacgaggtatcgccatcgcaaggacattgattcgaaatccaaagctg 3540  
DB 3481 tccggcggaacacgacgaggtatcgccatcgcaaggacattgattcgaaatccaaagctg 3540  
QY 3541 ttgtctacttgacgaagcgactcagccctcgacacgcaatcggaacgtcttgggttcaagct 3600  
DB 3541 ttgtctacttgacgaagcgactcagccctcgacacgcaatcggaacgtcttgggttcaagct 3600  
QY 3601 gccctcgatgaggcatccacgagccggaacacaaatagcagtggtgcgaccacttccact 3660  
DB 3601 gccctcgatgaggcatccacgagccggaacacaaatagcagtggtgcgaccacttccact 3660  
QY 3661 attcggaaattgagattatttttggttgttgcacaaaggagaatcgccgaaacgggcact 3720  
DB 3661 attcggaaattgagattatttttggttgttgcacaaaggagaatcgccgaaacgggcact 3720  
QY 3721 cgcgcggaactacaacgactgagaggaagatatattacgagatgtgttggcacaaatcttta 3780  
DB 3721 cgcgcggaactacaacgactgagaggaagatatattacgagatgtgttggcacaaatcttta 3780  
QY 3781 gaccaagcatga 3792  
DB 3781 gaccaagcatga 3792

## RESULT 5

US-09-677-488A-10  
; Sequence 10, Application US/09677488A  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; APPLICANT: Maddox, Joyce  
; APPLICANT: Gilliam, Jacob  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Crasta, Oswald R.  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin  
; FILE REFERENCE: 35718/204100  
; CURRENT APPLICATION NUMBER: US/09/677,488A  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12







|||||  
Db 1921 cgagaagctgctggfgaacgaccgacgtgagcgactcacaccactgccacatctcaa 1980  
QY 1981 gctggagacctggagaaagcgaagctgcggtcgcggaatttgggtactcgctcctaaaa 2040  
Db 1981 gctggagacctggagaaagcgaagctgcggtcgcggaatttgggtactcgctcctaaaa 2040  
QY 2041 tgcacctcaatcatgttctacgaacaaaaatctctactggtgctgttctgttgcaca 2100  
Db 2041 tgcacctcaatcatgttctacgaacaaaaatctctactggtgctgttctgttgcaca 2100  
QY 2101 ataacggttctgatgcgctgcacatttccagacaaagccctttgttttccgagattg 2160  
Db 2101 ataacggttctgatgcgctgcacatttccagacaaagccctttgttttccgagattg 2160  
QY 2161 ctcactgtcttcgagtgagtggtcatcgccacaggaacgggacagatcttttatcttg 2220  
Db 2161 ctcactgtcttcgagtgagtggtcatcgccacaggaacgggacagatcttttatcttg 2220  
QY 2221 atgttcttctgtcggtctaggaatctagtagatatttcacgattggctggacatgc 2280  
Db 2221 atgttcttctgtcggtctaggaatctagtagatatttcacgattggctggacatgc 2280  
QY 2281 aacgttatttcacaagttgtcaccatcgctatcaagcgcgaatttccaaacgagtaactg 2340  
Db 2281 aacgttatttcacaagttgtcaccatcgctatcaagcgcgaatttccaaacgagtaactg 2340  
QY 2341 gataacagactcgaaactcctcgaaatcccgagagaaatttctgtgctcacaatcgaa 2400  
Db 2341 gataacagactcgaaactcctcgaaatcccgagagaaatttctgtgctcacaatcgaa 2400  
QY 2401 ctgtcagctctcaccacgagttgcaagagttgatcatcagcaaatcttctcatttatc 2460  
Db 2401 ctgtcagctctcaccacgagttgcaagagttgatcatcagcaaatcttctcatttatc 2460  
QY 2461 gttctcggtcaactctgtctcgagcagtgctctaccactagcctatgtaggaactggg 2520  
Db 2461 gttctcggtcaactctgtctcgagcagtgctctaccactagcctatgtaggaactggg 2520  
QY 2521 cctggtggtgtgttgggtgaattccacccctgcttctgctggtcctcagaattcg 2580  
Db 2521 cctggtggtgtgttgggtgaattccacccctgcttctgctggtcctcagaattcg 2580  
QY 2581 tctagacgaagctagaagccggaactcggcaaaccttgcagaaagtctgggtctgc 2640  
Db 2581 tctagacgaagctagaagccggaactcggcaaaccttgcagaaagtctgggtctgc 2640  
QY 2641 aagcgaagcagttaccgcgatcccggaacgtctctatcttctgactctcgaagccatgtct 2700  
Db 2641 aagcgaagcagttaccgcgatcccggaacgtctctatcttctgactctcgaagccatgtct 2700  
QY 2701 ccaacagtaactcgaaatgttgacaaggcttctgaagatcaccaaaagcttttgggtt 2760  
Db 2701 ccaacagtaactcgaaatgttgacaaggcttctgaagatcaccaaaagcttttgggtt 2760  
QY 2761 ggaagatgttttgggttctcactgtcacagtcgatgaggttttgggtattgcccggga 2820  
Db 2761 ggaagatgttttgggttctcactgtcacagtcgatgaggttttgggtattgcccggga 2820  
QY 2821 ttttggattgcagtcgaataattgcttccagtgagtgagtcagcacacaactcaatttatatca 2880  
Db 2821 ttttggattgcagtcgaataattgcttccagtgagtgagtcagcacacaactcaatttatatca 2880  
QY 2881 tcttcgtggggttttgggttcccggtccaaagcagcagcccaagatttgggttactcaag 2940  
Db 2881 tcttcgtggggttttgggttcccggtccaaagcagcagcccaagatttgggttactcaag 2940  
QY 2941 agttttaccaaggtctggtcggtcgcaactatatactctggtcgcggaactgaagcgg 3000  
Db 2941 agttttaccaaggtctggtcggtcgcaactatatactctggtcggtcggaactgaagcgg 3000  
QY 3001 aacatccgcgaacgggagagaacaaaaagcccgagtggtgagtcctctgtcgac 3060  
|||||

Db 3001 accatccgcgaacgagagagaacaaagggccagggccagtggtggtggtggtccctgtcgac 3060  
QY 3061 ctcgagagacattgaattcaggttatcgtaacgtgattctcgctcgagtgcttcccgcggtt 3120  
Db 3061 ctcgagagacattgaattcaggttatcgtaacgtgattctcgctcgagtgcttcccgcggtt 3120  
QY 3121 tccatgacaatcgagcagcagacaattttagcttatgtggtcgcttctggtctggtgcaag 3180  
Db 3121 tccatgacaatcgagcagcagacaattttagcttatgtggtcgcttctggtctggtgcaag 3180  
QY 3181 tcaacgttgcactgttctgcgaacgattctacgacccgacctcggtcggaatttcaatt 3240  
Db 3181 tcaacgttgcactgttctgcgaacgattctacgacccgacctcggtcggaatttcaatt 3240  
QY 3241 gcaacgagaatatttcagaaaaatgtcgccgctgttaccgcggccatgtcttctggtc 3300  
Db 3241 gcaacgagaatatttcagaaaaatgtcgccgctgttaccgcggccatgtcttctggtc 3300  
QY 3301 caacaggaaccacacatttaccagagctcgttccgagagatttgcgagaaatgtgacctgaa 3360  
Db 3301 caacaggaaccacacatttaccagagctcgttccgagagatttgcgagaaatgtgacctgaa 3360  
QY 3361 gccgaattatcagaagagcttctcaagagcgtcttcccgcaagcccaatgcttggatt 3420  
Db 3361 gccgaattatcagaagagcttctcaagagcgtcttcccgcaagcccaatgcttggatt 3420  
QY 3421 ttgtcatcttcttaccagaaggtcttgaacgcttgcggtcctcaacgagggatgcagttc 3480  
Db 3421 ttgtcatcttcttaccagaaggtcttgaacgcttgcggtcctcaacgagggatgcagttc 3480  
QY 3481 tccgcggggaacacagcagcgagtcgcatcgccatcgcaagcattgattcgaaatccaaactg 3540  
Db 3481 tccgcggggaacacagcagcgagtcgcatcgccatcgcaagcattgattcgaaatccaaactg 3540  
QY 3541 ttgtcacttgcagaagcagctcagccctcgacacacaaatcggaagcttctggttcaagct 3600  
Db 3541 ttgtcacttgcagaagcagctcagccctcgacacacaaatcggaagcttctggttcaagct 3600  
QY 3601 gccctcgatgaggtatccacgagccgaacaaatagcagtggtgcacccagcttccact 3660  
Db 3601 gccctcgatgaggtatccacgagccgaacaaatagcagtggtgcacccagcttccact 3660  
QY 3661 attcggaattgattgatttatttcttgcgttgcacgggagaaatcgccgaacgggact 3720  
Db 3661 attcggaattgattgatttatttcttgcgttgcacgggagaaatcgccgaacgggact 3720  
QY 3721 cagcgggaactacaacgactgagagaaatattacgagatgtgttggcacatcttta 3780  
Db 3721 cagcgggaactacaacgactgagagaaatattacgagatgtgttggcacatcttta 3780  
QY 3781 gaccaagcatga 3792  
Db 3781 gaccaagcatga 3792

## RESULT 6

US-09-677-682-10  
; Sequence 10, Application US/09677682  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 35718/204101 (5718-111B)  
; CURRENT APPLICATION NUMBER: US/09/677,682  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 09/351,224  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 3792  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:

; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA  
us-09-677-682-10

Query Match 100.0%; Score 3792; DB 26; Length 3792;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgacagatgaatcggagaaacctcgacaaacccaagatggcagtgagctcctcacac 60  
Db 1 atgacagatgaatcggagaaacctcgacaaacccaagatggcagtgagctcctcacac 60  
Qy 61 cctcccagaaaagaaacgaaagcagatatttcagactatctcagaaatcttcagat 120  
Db 61 cctcccagaaaagaaacgaaagcagatatttcagactatctcagaaatcttcagat 120  
Qy 121 gccacaaatcacgactgactctcaatgcatcgcgctcatctgcccattggtccggg 180  
Db 121 gccacaaatcacgactgactctcaatgcatcgcgctcatctgcccattggtccggg 180  
Qy 181 gcttcctctctgactgactcactatcttcgtagcttcacaaagttcaacaattac 240  
Db 181 gcttcctctctgactgactcactatcttcgtagcttcacaaagttcaacaattac 240  
Qy 241 aattcggcgacggagtcctgaagcgttcaaggccgtagtgatcatttcgctctg 300  
Db 241 aattcggcgacggagtcctgaagcgttcaaggccgtagtgatcatttcgctctg 300  
Qy 301 ttgctctacctcttatttgggaagtgttgcctcacgtactgttccacggctgccattacc 360  
Db 301 ttgctctacctcttatttgggaagtgttgcctcacgtactgttccacggctgccattacc 360  
Qy 361 atttcagatatcaccaactcgaactcttcgacagtgcttccttgaaagcacttgagg 420  
Db 361 atttcagatatcaccaactcgaactcttcgacagtgcttccttgaaagcacttgagg 420  
Qy 421 caagggtctggcatttcgacaaagcagagcaatggagcaatgcgcactcargtcactacc 480  
Db 421 caagggtctggcatttcgacaaagcagagcaatggagcaatgcgcactcargtcactacc 480  
Qy 481 aatggcaaccgtatatacaaacaggtattgccgagaatattggtcttaccgtcaggcact 540  
Db 481 aatggcaaccgtatatacaaacaggtattgccgagaatattggtcttaccgtcaggcact 540  
Qy 541 tcaattgtctttctgactttgtgctgttggctcagtgagagtagcttttaac 600  
Db 541 tcaattgtctttctgactttgtgctgttggctcagtgagagtagcttttaac 600  
Qy 601 accatgtccgtcatccctgccattttcctgtgtccacggcattgtgcatagcaattgatgcc 660  
Db 601 accatgtccgtcatccctgccattttcctgtgtccacggcattgtgcatagcaattgatgcc 660  
Qy 661 gctcaggaggccaggatcaccaaggatctactacagcgcgcgtctcctcagaagaagtc 720  
Db 661 gctcaggaggccaggatcaccaaggatctactacagcgcgcgtctcctcagaagaagtc 720  
Qy 721 ttatcatccatccgacagatcaccaaggatctactacagcgcgcgtctcctcagaagaagtc 780  
Db 721 ttatcatccatccgacagatcaccaaggatctactacagcgcgcgtctcctcagaagaagtc 780  
Qy 781 gatgtcttttttcagcaagcaccacaaagagggaagaataatcgccaataatgsgtc 840  
Db 781 gatgtcttttttcagcaagcaccacaaagagggaagaataatcgccaataatgsgtc 840  
Qy 841 ttgttctcaactgagtaacttttgcatttcgctgctatcgactggccttttggaaggt 900  
Db 841 ttgttctcaactgagtaacttttgcatttcgctgctatcgactggccttttggaaggt 900  
Qy 901 ttctgcataatcagaatggcgaggttgcgcagcttggcaagctcttactgttgcctt 960  
Db 901 ttctgcataatcagaatggcgaggttgcgcagcttggcaagctcttactgttgcctt 960  
Qy 961 tccgtcaaccttagcagcaacgtccatctcaatgtcttgccctcaggttcagtcgtttac 1020

Db 961 tccgtcaaccttagcagcaacgtccatctcaatgtcttgccctcaggttcagtcgtttac 1020  
Qy 1021 caacgccgcatcttcgctccgaattattcagtatcatattgacaaacccacgcagctcgac 1080  
Db 1021 caacgccgcatcttcgctccgaattattcagtatcatattgacaaacccacgcagctcgac 1080  
Qy 1081 cctctgaccccttctggaagcagccagagggctgctaggtccaattgagatccaaaac 1140  
Db 1081 cctctgaccccttctggaagcagccagagggctgctaggtccaattgagatccaaaac 1140  
Qy 1141 ctggcaatttgcctaccctcccgaccatctgcccagttacttcgagatttcaactgaca 1200  
Db 1141 ctggcaatttgcctaccctcccgaccatctgcccagttacttcgagatttcaactgaca 1200  
Qy 1201 attccagctcggcaagcagcggccctcgctgctgcatcaggtagcgcaaaaagcacatg 1260  
Db 1201 attccagctcggcaagcagcggccctcgctgctgcatcaggtagcgcaaaaagcacatg 1260  
Qy 1261 gtcggcttacttgaaacgggtgtatctgcccagttcgggggaggtatttacttgatgggtg 1320  
Db 1261 gtcggcttacttgaaacgggtgtatctgcccagttcgggggaggtatttacttgatgggtg 1320  
Qy 1321 gaactgggacaatacaaatgtgaaatggctgagaagccgcatctcgcttcaacagaaa 1380  
Db 1321 gaactgggacaatacaaatgtgaaatggctgagaagccgcatctcgcttcaacagaaa 1380  
Qy 1381 cctgtgtgttcttcgtggcacaatctctccagaacatttccaaacgttttcattggtgagcaa 1440  
Db 1381 cctgtgtgttcttcgtggcacaatctctccagaacatttccaaacgttttcattggtgagcaa 1440  
Qy 1441 cgagatctgctcgcgaaacaaataatggagctgtgcaaaaagcttgcaaaagccagcaat 1500  
Db 1441 cgagatctgctcgcgaaacaaataatggagctgtgcaaaaagcttgcaaaagccagcaat 1500  
Qy 1501 ggcgacgtgtctaatatgacttcgcaacggttatgagactgaagtggcgagcgagcc 1560  
Db 1501 ggcgacgtgtctaatatgacttcgcaacggttatgagactgaagtggcgagcgagcc 1560  
Qy 1561 ggagccttgagtgaggtcaacgacaacgaattgcaatcgcacgaagtatcatatcgagat 1620  
Db 1561 ggagccttgagtgaggtcaacgacaacgaattgcaatcgcacgaagtatcatatcgagat 1620  
Qy 1621 cccaaagtcctgttactcgatgaagtaccagcgcccttgaccggaagcgaggaagtcg 1680  
Db 1621 cccaaagtcctgttactcgatgaagtaccagcgcccttgaccggaagcgaggaagtcg 1680  
Qy 1681 gtcagagggccttgaacggagtgctccaaagaccgactacttttggctatttggccacaaa 1740  
Db 1681 gtcagagggccttgaacggagtgctccaaagaccgactacttttggctatttggccacaaa 1740  
Qy 1741 ctgacactgtcaaaaagtgtcgcaacatcgagtcatttcccagggggaaaatcgctcgag 1800  
Db 1741 ctgacactgtcaaaaagtgtcgcaacatcgagtcatttcccagggggaaaatcgctcgag 1800  
Qy 1801 caaggcaacacacacgaattgatcgaaattcgctgctattacgcgcgactgggtcggtgca 1860  
Db 1801 caaggcaacacacacgaattgatcgaaattcgctgctattacgcgcgactgggtcggtgca 1860  
Qy 1861 caggacctcgggctgacgaacaacaagaacatgagaagaccctgcacgaagaagcgagca 1920  
Db 1861 caggacctcgggctgacgaacaacaagaacatgagaagaccctgcacgaagaagcgagca 1920  
Qy 1921 cgagaaagctgctggtgaacgcccgccacttgagcgactcacacacactgcacactctcaaa 1980  
Db 1921 cgagaaagctgctggtgaacgcccgccacttgagcgactcacacacactgcacactctcaaa 1980  
Qy 1981 gctggagaccttgagaagcggaaggtgccggtcgggactttgggctactctgctctaaaa 2040  
Db 1981 gctggagaccttgagaagcggaaggtgccggtcgggactttgggctactctgctctaaaa 2040  
Qy 2041 tgcacctaatactgattctacgaacaaaaaaatctctactggtgcttcttctgttgcacaa 2100







Query Match 100.0%; Score 3792; DB 26; Length 3792;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	atg	cag	at	g	a	a	a	g
Db	1	atg	cag	at	g	a	a	a	g
Qy	61	cct	ccc	cag	a	a	a	a	g
Db	61	cct	ccc	cag	a	a	a	a	g
Qy	121	g	cc	a	a	a	a	a	g
Db	121	g	cc	a	a	a	a	a	g
Qy	181	g	ctt	c	c	c	c	c	t
Db	181	g	ctt	c	c	c	c	c	t
Qy	241	a	at	t	c	c	c	c	a
Db	241	a	at	t	c	c	c	c	a
Qy	301	t	c	c	t	a	c	c	t
Db	301	t	c	c	t	a	c	c	t
Qy	361	a	t	t	c	a	c	c	a
Db	361	a	t	t	c	a	c	c	a
Qy	421	c	a	a	g	g	c	a	a
Db	421	c	a	a	g	g	c	a	a
Qy	481	a	a	t	t	c	c	a	a
Db	481	a	a	t	t	c	c	a	a
Qy	541	t	c	a	a	t	t	c	c
Db	541	t	c	a	a	t	t	c	c
Qy	601	a	c	a	t	t	c	c	a
Db	601	a	c	a	t	t	c	c	a
Qy	661	g	c	t	c	a	a	a	g
Db	661	g	c	t	c	a	a	a	g
Qy	721	t	a	t	c	a	a	a	t
Db	721	t	a	t	c	a	a	a	t
Qy	781	g	a	t	t	c	a	a	a
Db	781	g	a	t	t	c	a	a	a
Qy	841	t	t	t	c	a	a	a	g
Db	841	t	t	t	c	a	a	a	g
Qy	901	t	t	c	a	a	a	a	g
Db	901	t	t	c	a	a	a	a	g
Qy	961	t	c	c	a	a	a	a	g
Db	961	t	c	c	a	a	a	a	g
Qy	1021	c	a	a	c	c	c	a	a
Db	1021	c	a	a	c	c	c	a	a

```
Db 2101 ataaaggttctgatcgcgccacatttccaggagaagccctttgttttttcgagatcg 2160
QY 2161 ctcaactgtcttcgagttgaagtcatcgcgacaggaacacggcgacactttttatatctcg 2220
Db 2161 ctcaactgtcttcgagttgaagtcatcgcgacaggaacacggcgacactttttatatctcg 2220
QY 2221 atgttcttttgcgtggtctcagaaatctcagtagagatatattcacagatttgcctggacatcg 2280
Db 2221 atgttcttttgcgtggtctcagaaatctcagtagagatatattcacagatttgcctggacatcg 2280
QY 2281 aacgttatttccaaagtgtccaccatogctatcaagccgcaatgttccaaacgagtagctg 2340
Db 2281 aacgttatttccaaagtgtccaccatogctatcaagccgcaatgttccaaacgagtagctg 2340
QY 2341 gatcaagacatcgaactcctcgacatcccggaagcaaatcttctggtctctcacatcgcaa 2400
Db 2341 gatcaagacatcgaactcctcgacatcccggaagcaaatcttctggtctctcacatcgcaa 2400
QY 2401 ctgtcagctctaccacacgagttgcaagagttgatcagcaaatcttctcatttatatc 2460
Db 2401 ctgtcagctctaccacacgagttgcaagagttgatcagcaaatcttctcatttatatc 2460
QY 2461 gttgcggtcacaactcgtctcagcagtgctctcacactagcctatggaatggaactggg 2520
Db 2461 gttgcggtcacaactcgtctcagcagtgctctcacactagcctatggaatggaactggg 2520
QY 2521 cctgggtgtgtgtgtgtgcacttccacctgcttctgttggctggtcactcagaattcg 2580
Db 2521 cctgggtgtgtgtgtgtgcacttccacctgcttctgttggctggtcactcagaattcg 2580
QY 2581 tctagagacgaagctagaagccggaaactcggcaactttgcagaaagtgcctgggtctgc 2640
Db 2581 tctagagacgaagctagaagccggaaactcggcaactttgcagaaagtgcctgggtctgc 2640
QY 2641 aagcgaacgagttacccgcagtcgcagcagtgctctcactagcctatggaatggaactggg 2700
Db 2641 aagcgaacgagttacccgcagtcgcagcagtgctctcactagcctatggaatggaactggg 2700
QY 2701 ccaacagtaactcgcagatgttgagcaaggtcttctgaagatcactcgaagcttttggttt 2760
Db 2701 ccaacagtaactcgcagatgttgagcaaggtcttctgaagatcactcgaagcttttggttt 2760
QY 2761 ggaagatgttttgggttttcaactgttcaacagtcagtgaggttttggctattgcccctggga 2820
Db 2761 ggaagatgttttgggttttcaactgttcaacagtcagtgaggttttggctattgcccctggga 2820
QY 2821 ttttgcattgcagtcgataaattggttcaggtcagtgagtagcagacacactcaattttatca 2880
Db 2821 ttttgcattgcagtcgataaattggttcaggtcagtgagtagcagacacactcaattttatca 2880
QY 2881 tcttcgtggggttttgggttcgggtcgaactatctctggtcgcgacatttgcttaactccacg 2940
Db 2881 tcttcgtggggttttgggttcgggtcgaactatctctggtcgcgacatttgcttaactccacg 2940
QY 2941 agttttaccaaggtcgtggttcgggtcgaactatctctggtcgcgacatttgcttaactccacg 3000
Db 2941 agttttaccaaggtcgtggttcgggtcgaactatctctggtcgcgacatttgcttaactccacg 3000
QY 3001 accatccgcgaacacggaggaacaaagaaagcccgagtggtgagtcgctgtcgac 3060
Db 3001 accatccgcgaacacggaggaacaaagaaagcccgagtggtgagtcgctgtcgac 3060
QY 3061 ctcgagacattgaattcagtagtcgtaacgtagtcgctcagcagccagtagtttgcttaactccacg 3120
Db 3061 ctcgagacattgaattcagtagtcgtaacgtagtcgctcagcagccagtagtttgcttaactccacg 3120
QY 3121 tccatgacaatcgcagccaggacaatttgcattatgtggtcgtcttctggtgtggaag 3180
Db 3121 tccatgacaatcgcagccaggacaatttgcattatgtggtcgtcttctggtgtggaag 3180
QY 3181 tcaacgttgatcgtttgtcgggaacgattctacgaccgcacctcgggcgaatttcattt 3240
Db 3181 tcaacgttgatcgtttgtcgggaacgattctacgaccgcacctcgggcgaatttcattt 3240
```

```
QY 3241 gcacacagagaatatgttcagaaatgtccgcgcttgcacgcggccatattgtctttggtc 3300
Db 3241 gcacacagagaatatgttcagaaatgtccgcgcttgcacgcggccatattgtctttggtc 3300
QY 3301 caacaggaacccacactttaccaaaggtcgttcgcgagaaatgtgaagtgtggcctcgaa 3360
Db 3301 caacaggaacccacactttaccaaaggtcgttcgcgagaaatgtgaagtgtggcctcgaa 3360
QY 3361 gccgaattatcagaagagctttgtcaaggagcgttcccgcaagcgaatcttttgatt 3420
Db 3361 gccgaattatcagaagagctttgtcaaggagcgttcccgcaagcgaatcttttgatt 3420
QY 3421 ttgtcatctctttaccagaaggtttgaaacgcttgcggtcgaagcaggttcgagttc 3480
Db 3421 ttgtcatctctttaccagaaggtttgaaacgcttgcggtcgaagcaggttcgagttc 3480
QY 3481 tccggcgggcaacgacgagcgttcgcatcgcaagagcattgattcgaatccaaagctg 3540
Db 3481 tccggcgggcaacgacgagcgttcgcatcgcaagagcattgattcgaatccaaagctg 3540
QY 3541 ttgtcactttgacgaagcagctcagccctcgacgcgaatcggaacgtctgtgttcaagct 3600
Db 3541 ttgtcactttgacgaagcagctcagccctcgacgcgaatcggaacgtctgtgttcaagct 3600
QY 3601 gccctcgatgagcagcagcagcccgcaacacacacacacacacacacacacacacac 3660
Db 3601 gccctcgatgagcagcagcagcccgcaacacacacacacacacacacacacacacac 3660
QY 3661 attcggaaatgtgatgttattttgttgcgaacggagaaatcgccgaacgggagcact 3720
Db 3661 attcggaaatgtgatgttattttgttgcgaacgggagaaatcgccgaacgggagcact 3720
QY 3721 cagcgggaactacaacgactgaggaagatatattacgagatgtgtttggcacaatcttta 3780
Db 3721 cagcgggaactacaacgactgaggaagatatattacgagatgtgtttggcacaatcttta 3780
QY 3781 gaccaagcatga 3792
Db 3781 gaccaagcatga 3792
```

## RESULT 9

```
US-09-882-694-10
; Sequence 10, Application US/09882694
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 35718/208255
; CURRENT APPLICATION NUMBER: US/09/882,694
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 3792
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: p-glycoprotein, fully spliced cDNA
US-09-882-694-10
```

```
Query Match 100.0%; Score 3792; DB 33; Length 3792;
Best Local Similarity 100.0%; Pred. No. 0;
```



Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	atggcagatgaatcggagaaacctcgacacaaaccaaagtggcagtgagtcgtctctcaac	60	
Db	1	atggcagatgaatcggagaaacctcgacacaaaccaaagtggcagtgagtcgtctctcaac	60	
Qy	61	ctctcccccagaaagaaacgaagcagatattcagactactctcgaatcttcagatat	120	
Db	61	ctctcccccagaaagaaacgaagcagatattcagactactctcgaatcttcagatat	120	
Qy	121	gccgacaatacagctgactctcaatgtcatcgctcatctgcgccatcgatccggg	180	
Db	121	gccgacaatacagctgactctcaatgtcatcgctcatctgcgccatcgatccggg	180	
Qy	181	gttcccttctctgatgtcgatcatctcttcggtagcttcacacaaagaattcaacaattac	240	
Db	181	gttcccttctctgatgtcgatcatctcttcggtagcttcacacaaagaattcaacaattac	240	
Qy	241	aattcggagcagcgagtcctgaagcgttcaaggccgatgtggatcatttcgtccctgtgg	300	
Db	241	aattcgggagcagcgagtcctgaagcgttcaaggccgatgtggatcatttcgtccctgtgg	300	
Qy	301	ttcgtctacacctttattggggaagttgtcctcaagtaagtttccacggctgccattacc	360	
Db	301	ttcgtctacacctttattggggaagttgtcctcaagtaagtttccacggctgccattacc	360	
Qy	361	atttcagctatcacgaacctcgaaactcttcgacgagttcttgaatgcaccttcggg	420	
Db	361	atttcagctatcacgaacctcgaaactcttcgacgagttcttctggaatgcaccttcggg	420	
Qy	421	caagaggtctgcatcttcgacaaagcagacaatggagcaatcgccactcargtcaactacc	480	
Db	421	caagaggtctgcatcttcgacaaagcagacaatggagcaatcgccactcargtcaactacc	480	
Qy	481	aatgcaacccgtatatacaaacaggtattgcgagaaatttggtcttcacgtgcaggcaatt	540	
Db	481	aatggaacccgtatatacaaacaggtattgcgagaaatttggtcttcacgtgcaggcaatt	540	
Qy	541	tcaatgtctttcttgcatcttgctgcttttgccgtctcagtggaagctagctttaac	600	
Db	541	tcaatgtctttcttgcatcttgctgcttttgccgtctcagtggaagctagctttaac	600	
Qy	601	accatgtccgtcatccctgccattttccctggtcacgcgcatctgcataattgatgcc	660	
Db	601	accatgtccgtcatccctgccattttccctggtcacgcgcatctgcataattgatgcc	660	
Qy	661	gctcaggagccaggaataccaggaatactcaacgcgcgctgtctcgcagagaagtc	720	
Db	661	gctcaggagccaggaataccaggaatactcaacgcgcgctgtctcgcagagaagtc	720	
Qy	721	ttatcatccatccggacagtcctatgctttctacgcccagaagaaatggtcgaaaaatat	780	
Db	721	ttatcatccatccggacagtcctatgctttctacgcccagaagaaatggtcgaaaaatat	780	
Qy	781	gatgtctttttgagcaagcacaccacaaagggaagaaagaaatcgccaaataatgsgtc	840	
Db	781	gatgtctttttgagcaagcacaccacaaagggaagaaagaaatcgccaaataatgsgtc	840	
Qy	841	ttgttctcaactgagtaatttgcattacgtctgtatcgactgccttttggaaggt	900	
Db	841	ttgttctcaactgagtaatttgcattacgtctgtatcgactgccttttggaaggt	900	
Qy	901	tttcgcatgtatcagaatggcgaggttgcgacgtttggcaagctcttactgtgtcctt	960	
Db	901	tttcgcatgtatcagaatggcgaggttgcgacgtttggcaagctcttactgtgtcctt	960	
Qy	961	tcgctcaacttagcagccacgtccatctcaatgtcttgcccttcaggttcagtcgtttac	1020	
Db	961	tcgctcaacttagcagccacgtccatctcaatgtcttgcccttcaggttcagtcgtttac	1020	
Qy	1021	caacgcgcactctcgctccgataattatcagtatcatcagaaacccacgcagctcgac	1080	
Db	1021	caacgcgcactctcgctccgataattatcagtatcatcagaaacccacgcagctcgac	1080	

Qy	1081	ctctcgaccctctctggaagcagccagagggctcctagatcaaatggatccaaaac	1140	
Db	1081	ctctcgaccctctctggaagcagccagagggctcctagatcaaatggatccaaaac	1140	
Qy	1141	ctggcatttgctaccctcccccaccatctgcccnaagtaacttcgagattcaacttgaca	1200	
Db	1141	ctggcatttgctaccctcccccaccatctgcccnaagtaacttcgagattcaacttgaca	1200	
Qy	1201	attccagctggcagaagcagcgccctctggtgctcatcaggtagcggcaaaagcacaatg	1260	
Db	1201	attccagctggcagaagcagcgccctctggtgctcatcaggtagcggcaaaagcacaatg	1260	
Qy	1261	gtcggcttacttgaaacggtggtatctcccaagttcggggaggaataattacttgatgggtg	1320	
Db	1261	gtcggcttacttgaaacggtggtatctcccaagttcggggaggaataattacttgatgggtg	1320	
Qy	1321	gaactgggacaatacaatgtgaatggctgagaaagcgcatctccctcgttcaacagaaa	1380	
Db	1321	gaactgggacaatacaatgtgaatggctgagaaagcgcatctccctcgttcaacagaaa	1380	
Qy	1381	cctgtgtgttctggtggcacaaatctccagaacattgccaaacgtttccatggatgagcaa	1440	
Db	1381	cctgtgtgttctggtggcacaaatctccagaacattgccaaacgtttccatggatgagcaa	1440	
Qy	1441	cgagatctgctcgcgnaaaacaaaatggagcttctgcaaaaagcttgcagaagccagcaat	1500	
Db	1441	cgagatctgctcgcgnaaaacaaaatggagcttctgcaaaaagcttgcagaagccagcaat	1500	
Qy	1501	ggcagcgttcttaattatgagcttccgaaacggttatgagactgaagttggcgagcgagcc	1560	
Db	1501	ggcagcgttcttaattatgagcttccgaaacggttatgagactgaagttggcgagcgagcc	1560	
Qy	1561	ggagccttgagtgaggttcaacgcaacgaattgcaatcgcaacgaagtatacatacggat	1620	
Db	1561	ggagccttgagtgaggttcaacgcaacgaattgcaatcgcaacgaagtatacatacggat	1620	
Qy	1621	cccaagatcctgttactcgtatgaagctaccagcgcccttgaccgaaagggcgagaaagtg	1680	
Db	1621	cccaagatcctgttactcgtatgaagctaccagcgcccttgaccgaaagggcgagaaagtg	1680	
Qy	1681	gtccagaggccttgaaacggagtgccaaagaccgcactacttttggtcatctgcccacaaa	1740	
Db	1681	gtccagaggccttgaaacggagtgccaaagaccgcactacttttggtcatctgcccacaaa	1740	
Qy	1741	ctagccactgtcaaaagtgtcggcaacatcgagtcatttcccagggggaaaaatcgctgag	1800	
Db	1741	ctagccactgtcaaaagtgtcggcaacatcgagtcatttcccagggggaaaaatcgctgag	1800	
Qy	1801	caaggcacacaccacgaattgatcgaattcggctgtcatattacgcgcgactggtgctgca	1860	
Db	1801	caaggcacacaccacgaattgatcgaattcggctgtcatattacgcgcgactggtgctgca	1860	
Qy	1861	caggacactcgggctgcagaaacaaacgaatacagagaccctgcagaaaggcgagca	1920	
Db	1861	caggacactcgggctgcagaaacaaacgaatacagagaccctgcagaaaggcgagca	1920	
Qy	1921	cgagaagctgctggtgaacgaccgcgacttgagcgactcacaccactgccacatatcaaa	1980	
Db	1921	cgagaagctgctggtgaacgaccgcgacttgagcgactcacaccactgccacatatcaaa	1980	
Qy	1981	gctggagacacttgagaaagcggaaggtgcgggtcgggactttgggtactcgcctccataaa	2040	
Db	1981	gctggagacacttgagaaagcggaaggtgcgggtcgggactttgggtactcgcctccataaa	2040	
Qy	2041	tgcacacttaactgttctacgaacaaaaaatctctactggtgcttcttctgttgtcaaca	2100	
Db	2041	tgcacacttaactgttctacgaacaaaaaatctctactggtgcttcttctgttgtcaaca	2100	
Qy	2101	ataacggttcttgatgtgcggcgccacatttccagagaaagccctttgttttcgagattcg	2160	
Db	2101	ataacggttcttgatgtgcggcgccacatttccagagaaagccctttgttttcgagattcg	2160	





QY 1 atgcagatgaatcgagaaacctcgaccacaaacaaagatgcagtgagtcgtccacac 60  
Db 1 atgagatgaatcgagaaacctcgaccacaaacaaagatgcagtgagtcgtccacac 60  
QY 61 cctccccagaaaaagaaacgaaggagtagtatttcagactactacgaatcttcagatat 120  
Db 61 cctccccagaaaaagaaacgaaggagtagtatttcagactactacgaatcttcagatat 120  
QY 121 gccgacaaatacagactggacttcaatgcatcgcgctcaatctgagccatcggaatccggg 180  
Db 121 gccgacaaatacagactggacttcaatgcatcgcgctcaatctgagccatcggaatccggg 180  
QY 181 gctccctctctcgatgcagatcatctcggtagcttcacacaaagatcaacaattac 240  
Db 181 gctccctctctcgatgcagatcatctcggtagcttcacacaaagatcaacaattac 240  
QY 241 aattcggggcagcgagtcctgaagcgttcaaggccgatgfggatacttcgtctgtgg 300  
Db 241 aattcggggcagcgagtcctgaagcgttcaaggccgatgfggatacttcgtctgtgg 300  
QY 301 ttctctacactcttattgggaagttagtctcaagtcaggtttccacggctgccaattacc 360  
Db 301 ttctctacactcttattgggaagttagtctcaagtcaggtttccacggctgccaattacc 360  
QY 361 atttcagctatacgaaccactcgaaactcttcgacgagtggttcttgaatgcacctgggg 420  
Db 361 atttcagctatacgaaccactcgaaactcttcgacgagtggttcttgaatgcacctgggg 420  
QY 421 caagaggctctggcatttcgacaagcagagcaaatggagcaaatcgccactcargtcaactacc 480  
Db 421 caagaggctctggcatttcgacaagcagagcaaatggagcaaatcgccactcargtcaactacc 480  
QY 481 aatggcaacctatatacaaacaggtattcccgagaaatggctttaccgtgcagcacctt 540  
Db 481 aatggcaacctatatacaaacaggtattcccgagaaatggctttaccgtgcagcacctt 540  
QY 541 tcaatgtctctttctgctattgtgctgctttggcgctcagtggaagtagctttaatc 600  
Db 541 tcaatgtctctttctgctattgtgctgctttggcgctcagtggaagtagctttaatc 600  
QY 601 accatgtctcgatccctggcaatttctctgtgtcaacggcatctgcatagcaattgatgcc 660  
Db 601 accatgtctcgatccctggcaatttctctgtgtcaacggcatctgcatagcaattgatgcc 660  
QY 661 gctcagaggccagatcaccagagatctactcacgcccgtctgtctcgacagaagaagtc 720  
Db 661 gctcagaggccagatcaccagagatctactcacgcccgtctgtctcgacagaagaagtc 720  
QY 721 ttatcatcatccggacagtcctatcttctacgcccagaaagaaatggtcgaaaaatat 780  
Db 721 ttatcatcatccggacagtcctatcttctacgcccagaaagaaatggtcgaaaaatat 780  
QY 781 gatgctcttttgcagcaagcacacaagaagggaagaaatcgccaaataatggsgtgc 840  
Db 781 gatgctcttttgcagcaagcacacaagaagggaagaaatcgccaaataatggsgtgc 840  
QY 841 ttgttctcaactgagtagcttttgcatctacgtctgtatcgactggcccttttggaaagt 900  
Db 841 ttgttctcaactgagtagcttttgcatctacgtctgtatcgactggcccttttggaaagt 900  
QY 901 ttctcgatgtatcagaatggcgaggttgccgacgttggcgaagtcttactgttcttact 960  
Db 901 ttctcgatgtatcagaatggcgaggttgccgacgttggcgaagtcttactgttcttact 960  
QY 961 tccgtcaaccttagagcaacgttcaatctcaatgcttggccttcaggttcagctgtttac 1020  
Db 961 tccgtcaaccttagagcaacgttcaatctcaatgcttggccttcaggttcagctgtttac 1020  
QY 1021 caacgccgatcttcggctccgaattattcagtagtataatggacaaaccccgagctcgac 1080  
Db 1021 caacgccgatcttcggctccgaattattcagtagtataatggacaaaccccgagctcgac 1080  
QY 1081 cctctcgacccttcttgaaaagcagccagaggggctgctaggttcaaatggagatccaaaaac 1140

Db 1081 cctctcgacccttcttgaaaagcagccagagggctgctaggtcaaatggagatccaaaac 1140  
QY 1141 ctggcatttgctacacccctcccgaccatctgcccagatcttcagatttcaacttgaca 1200  
Db 1141 ctggcatttgctacacccctcccgaccatctgcccagatcttcagatttcaacttgaca 1200  
QY 1201 attcagactggcagacgagccctcgctcggtgcatcaggtagcggaacaaagcacaaatg 1260  
Db 1201 attcagactggcagacgagccctcgctcggtgcatcaggtagcggaacaaagcacaaatg 1260  
QY 1261 gtcggcttacttgaaacggttggtatctgcccagttccggggaggatattactgtatgggttg 1320  
Db 1261 gtcggcttacttgaaacggttggtatctgcccagttccggggaggatattactgtatgggttg 1320  
QY 1321 gaactggggaacaatacaatgtgaaatggctgagagccgacttcgctcgtttcaacagaaa 1380  
Db 1321 gaactggggaacaatacaatgtgaaatggctgagagccgacttcgctcgtttcaacagaaa 1380  
QY 1381 cctgtgttcttcgtggcacaatcttccagaaacttgcgaacggtttcatggatgagcaa 1440  
Db 1381 cctgtgttcttcgtggcacaatcttccagaaacttgcgaacggtttcatggatgagcaa 1440  
QY 1441 cgagatctgcctcgcaaaaaaactggagcttgcgaacaaagcttgcgaacgagcaat 1500  
Db 1441 cgagatctgcctcgcaaaaaaactggagcttgcgaacaaagcttgcgaacgagcaat 1500  
QY 1501 ggcgacgttcttcaatgaagcttcggaacggttatagagactgaagtggcgagcgagcc 1560  
Db 1501 ggcgacgttcttcaatgaagcttcggaacggttatagagactgaagtggcgagcgagcc 1560  
QY 1561 ggagccttgagtgagggtcaacgacaaacttgcgaatcgcaagaaatatactatcgat 1620  
Db 1561 ggagccttgagtgagggtcaacgacaaacttgcgaatcgcaagaaatatactatcgat 1620  
QY 1621 cccaaagatcctgttactcgatgaagcttaccagggcccttgcacccgaaggcggaagatg 1680  
Db 1621 cccaaagatcctgttactcgatgaagcttaccagggcccttgcacccgaaggcggaagatg 1680  
QY 1681 gtcagagggccttgaaacgagtgctccaaagacgcgactacttggctcattggccacaaa 1740  
Db 1681 gtcagagggccttgaaacgagtgctccaaagacgcgactacttggctcattggccacaaa 1740  
QY 1741 ctaggcactgtcaaaaagtgtcggcaacatcgagtcatttcccaggggaaaaatcgctcag 1800  
Db 1741 ctaggcactgtcaaaaagtgtcggcaacatcgagtcatttcccaggggaaaaatcgctcag 1800  
QY 1801 caaggcacacacacagaaattgacgaattcggtctgctattacgcgcgactggtgctgca 1860  
Db 1801 caaggcacacacacagaaattgacgaattcggtctgctattacgcgcgactggtgctgca 1860  
QY 1861 caggacctcggggctgacgaacaaacaaagacatgagaaagcccttgcagaaaaaggcagca 1920  
Db 1861 caggacctcggggctgacgaacaaacaaagacatgagaaagcccttgcagaaaaaggcagca 1920  
QY 1921 cgagaagctgctggtgaaacgacccggaacttgcagcgcactcaacacactgcacacttcaa 1980  
Db 1921 cgagaagctgctggtgaaacgacccggaacttgcagcgcactcaacacactgcacacttcaa 1980  
QY 1981 gctggagacctggagaagcggaaggtgcggctcgggactttgggctactcgtctcctaaaa 2040  
Db 1981 gctggagacctggagaagcggaaggtgcggctcgggactttgggctactcgtctcctaaaa 2040  
QY 2041 tgcattcctaatactgttctacgaacaaaaaaatctctacttggtgcttcttctgttcaaca 2100  
Db 2041 tgcattcctaatactgttctacgaacaaaaaaatctctacttggtgcttcttctgttcaaca 2100  
QY 2101 ataacggttctgatctgcggccacatttccagagaacagccctttgttttctcgagattg 2160  
Db 2101 ataacggttctgatctgcggccacatttccagagaacagccctttgttttctcgagattg 2160  
QY 2161 ctactgctctcaggttgatggttcacgcgcagacagggcgagacttttatattctg 2220

Db 2161 ctcaactgtcttcgagttgagtggtcatgcgacaggaacggcgacacactttttatatcttg 2220  
QY 2221 atgtcttttgtgtggtctagaaatctagtaggatatattcaogattggtcggacatgc 2280  
Db 2221 atgtcttttgtgtggtctagaaatctagtaggatatattcaogattggtcggacatgc 2280  
QY 2281 aacgttatttcaaaagtgtccaccatcgctatcaacgcgcgaatgttccaaagagtaactg 2340  
Db 2281 aacgttatttcaaaagtgtccaccatcgctatcaacgcgcgaatgttccaaagagtaactg 2340  
QY 2341 gatcaagacatgaactctctcgacatcccgagagcaaatcttctggtctctacatcgcaa 2400  
Db 2341 gatcaagacatgaactctctcgacatcccgagagcaaatcttctggtctctacatcgcaa 2400  
QY 2401 ctgtcagctctaccacgcagttgcaagagttgatataogcaaaattttctcatttatatc 2460  
Db 2401 ctgtcagctctaccacgcagttgcaagagttgatataogcaaaattttctcatttatatc 2460  
QY 2461 gttgtcgggtcaacatcgctctcagcagtgctctaccactagcctatgagatgaaactggg 2520  
Db 2461 gttgtcgggtcaacatcgctctcagcagtgctctaccactagcctatgagatgaaactggg 2520  
QY 2521 cctgggtgtgtgttgggtgaactccacccctgcttttgggtggtcactcagaaattcg 2580  
Db 2521 cctgggtgtgtgttgggtgaactccacccctgcttttgggtggtcactcagaaattcg 2580  
QY 2581 tctagagcgaagctagaaagccggaactcggcaaaatttgcaagaagtgtcgggttgc 2640  
Db 2581 tctagagcgaagctagaaagccggaactcggcaaaatttgcaagaagtgtcgggttgc 2640  
QY 2641 aagcgaagcagttaccgcgactccgcgacgtctcatctttgactctcgaaggccatgttct 2700  
Db 2641 aagcgaagcagttaccgcgactccgcgacgtctcatctttgactctcgaaggccatgttct 2700  
QY 2701 ccaacagtaactcgcagactgttgacaaggtcttgtaagatactcaaaagcttttgggtt 2760  
Db 2701 ccaacagtaactcgcagactgttgacaaggtcttgtaagatactcaaaagcttttgggtt 2760  
QY 2761 ggaagatgtttgttttcaactgtccagtcgagtgagttgttttggctatgcccgtgga 2820  
Db 2761 ggaagatgtttgttttcaactgtccagtcgagtgagttgttttggctatgcccgtgga 2820  
QY 2821 tttgtattgcagtcgataattggtctcagtgagtagcagacacaaactcaattttatca 2880  
Db 2821 tttgtattgcagtcgataattggtctcagtgagtagcagacacaaactcaattttatca 2880  
QY 2881 tcttcgtggggttttgttccgggtccaaagcagcagcccagtatgttgcttaactccacg 2940  
Db 2881 tcttcgtggggttttgttccgggtccaaagcagcagcccagtatgttgcttaactccacg 2940  
QY 2941 agttttaccaggtcgtggtcgtgcgaactatatactctgctgcgcgacattgaagccg 3000  
Db 2941 agttttaccaggtcgtggtcgtgcgaactatatactctgctgcgcgacattgaagccg 3000  
QY 3001 accatccgcgaacgaggaggaacaaagaaagcccgatgggtggatgcccgtgac 3060  
Db 3001 accatccgcgaacgaggaggaacaaagaaagcccgatgggtggatgcccgtgac 3060  
QY 3061 ctgcagacattgaattcaggtatcgtcaacgtgattcgtgcgtcagttctcccggggtt 3120  
Db 3061 ctgcagacattgaattcaggtatcgtcaacgtgattcgtgcgtcagttctcccggggtt 3120  
QY 3121 tccatgacaatcgcagcaggaacaattttagcttatgtgggcgctctctgctgtggaag 3180  
Db 3121 tccatgacaatcgcagcaggaacaattttagcttatgtgggcgctctctgctgtggaag 3180  
QY 3181 tcaacgttgatcgtcttgcggaaacgattctacacccgacctcggccgaatttcattt 3240  
Db 3181 tcaacgttgatcgtcttgcggaaacgattctacacccgacctcggccgaatttcattt 3240  
QY 3241 gcacacgagaatattgcagaaatgtcgcggttgcacgcggcccatatgtctttggtc 3300  
Db 3241 gcacacgagaatattgcagaaatgtcgcggttgcacgcggcccatatgtctttggtc 3300

QY 3301 caacaggaacccacacatttaccagggtccqcttcgcgagaatgtgacgttgccctcgaa 3360  
Db 3301 caacaggaacccacacatttaccagggtccqcttcgcgagaatgtgacgttgccctcgaa 3360  
QY 3361 gcggaattatcagaagagctttgtcaaggagcgccttcccgcaagccaatgtcttggatt 3420  
Db 3361 gcggaattatcagaagagctttgtcaaggagcgccttcccgcaagccaatgtcttggatt 3420  
QY 3421 ttgtcatctctttaccagaaggtcttgaaacgccttcggtcctcaacagggatcagttc 3480  
Db 3421 ttgtcatctctttaccagaaggtcttgaaacgccttcggtcctcaacagggatcagttc 3480  
QY 3481 tccggcgggcaacagcagcggatcgcctatcgcaagagcattgattcgaatcccaagctg 3540  
Db 3481 tccggcgggcaacagcagcggatcgcctatcgcaagagcattgattcgaatcccaagctg 3540  
QY 3541 ttgtacttgacaagagcagctcagccctcgacacgcaatacggaaactgtctggttcaagct 3600  
Db 3541 ttgtacttgacaagagcagctcagccctcgacacgcaatacggaaactgtctggttcaagct 3600  
QY 3601 gcctcgtatgaggtcctccacgagccgaacgacaatagcagtgcgccagcactttccact 3660  
Db 3601 gcctcgtatgaggtcctccacgagccgaacgacaatagcagtgcgccagcactttccact 3660  
QY 3661 attcggaaattgtgatgttattttgtgtttgccaaaggagaatcgcgaaacgggcact 3720  
Db 3661 attcggaaattgtgatgttattttgtgtttgccaaaggagaatcgcgaaacgggcact 3720  
QY 3721 cagcggaactacaacgactcagagagaagataattacagatgtgtttggcacaattctta 3780  
Db 3721 cagcggaactacaacgactcagagagaagataattacagatgtgtttggcacaattctta 3780  
QY 3781 gaccaagcatga 3792  
Db 3781 gaccaagcatga 3792

RESULT 11

US-09-351-224-9  
; Sequence 9, Application US/09351224  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jon  
; TITLE OF INVENTION: Compositions and Methods for Fummonisin Detoxification  
; FILE REFERENCE: 5718-111  
; CURRENT APPLICATION NUMBER: US/09/351,224  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 3999  
; TYPE: DNA  
; ORGANISM: Exophiala spinifera  
; FEATURE:  
; OTHER INFORMATION: p-glycoprotein, with introns  
US-09-351-224-9

Query Match 20.6%; Score 780; DB 17; Length 3999;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 830; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggcagatgaatcggagaaaactcgcacaaacccaagatggcgatgagtcgtctccacac 60  
Db 20 atggcagatgaatcggagaaaactcgcacaaacccaagatggcgatgagtcgtctccacac 79  
QY 61 cctccccagaaaaagaaacccaagcagattttcagactatctacgaattcttcagatat 120  
Db 80 cctccccagaaaaagaaacccaagcagattttcagactatctacgaattcttcagatat 139  
QY 121 gccgacaatacgcactggactctcaatgtcactcgcgtctcatctgcgccatcgatccggg 180  
Db 140 gccgacaatacgcactggactctcaatgtcactcgcgtctcatctgcgccatcgatccggg 199

```
QY 181 gcttcctctctctgtagtgcagatcattctggttagcttccaccacaaagttcaacaattac 240
Db 200 gcttcctctctctgtagtgcagatcattctggttagcttccaccacaaagttcaacaattac 259
QY 241 aattcgggacagggagctcctgaagcgttcaagccgagatgagatcattctcctctgtg 300
Db 260 aattcgggacagggagctcctgaagcgttcaagccgagatgagatcattctcctctgtg 319
QY 301 ttggttacctcttatttgggaagtgttctctcaacgtacgtatttccacgggtgcattacc 360
Db 320 ttggttacctcttatttgggaagtgttctctcaacgtacgtatttccacgggtgcattacc 379
QY 361 atttcagctataagaaacactcgaactcttcgcagcaggtgttcttgaatgcaccttgcgg 420
Db 380 atttcagctataagaaacactcgaactcttcgcagcaggtgttcttgaatgcaccttgcgg 439
QY 421 caagaggtctggcatttcgacaagcagagcaatggagcaatcgccactcargtcaactacc 480
Db 440 caagaggtctggcatttcgacaagcagagcaatggagcaatcgccactcargtcaactacc 499
QY 481 aatggcaacgctatacaaacaggtattgcgagaaatttggcttaccgtgcaggcaactt 540
Db 500 aatggcaacgctatacaaacaggtattgcgagaaatttggcttaccgtgcaggcaactt 559
QY 541 tcaatgtctcttctgcatttggctgcttggcgtctcagtggaagctagctttaaacc 600
Db 560 tcaatgtctcttctgcatttggctgcttggcgtctcagtggaagctagctttaaacc 619
QY 601 accatgtccgtatccctgcatttctctggttcacccggcatctgcataattgagtc 660
Db 620 accatgtccgtatccctgcatttctggttcacccggcatctgcataattgagtc 679
QY 661 gctcagagccagagatcaccaagatctactacgcgcgcgtctctcgcagaagaagtc 720
Db 680 gctcagagccagagatcaccaagatctactacgcgcgcgtctctcgcagaagaagtc 739
QY 721 ttatcatcatcgcgacagtccttctctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
Db 740 ttatcatcatcgcgacagtccttctctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 799
QY 781 gatgtcttttgcagcaagcacacacaaaggaagggaagaaatcgccaaat 831
Db 800 gatgtcttttgcagcaagcacacacaaaggaagggaagaaatcgccaaat 850
```

## RESULT 12

```
US-09-351-224E-9
; Sequence 9, Application US/09351224E
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 5718-111
; CURRENT APPLICATION NUMBER: US/09/351,224E
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: p-glycoprotein, with introns
US-09-351-224E-9
```

```
Query Match 20.6%; Score 780; DB 17; Length 3999;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 830; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgacagatgaatcgagaaacctcgaccacaaacagatggcagtgagtcgtcctccacac 60
Db 20 atggcagatgaatcgagaaacctcgaccacaaacagatggcagtgagtcgtcctccacac 79
QY 61 cctcccccgaagaaagaaacccgaagcagatttccagactatctacgaattctcagatat 120
Db 80 cctcccccgaagaaagaaacccgaagcagatttccagactatctacgaattctcagatat 139
QY 121 gcgcacaaatacagactgggaacttcaatgtcactgcgtcactctgcgccatcgagtcggg 180
Db 140 gcgcacaaatacagactgggaacttcaatgtcactgcgtcactctgcgccatcgagtcggg 199
QY 181 gcttcctctctctgtagtgcacatcttctggttagcttccacaaacagtttcaacaattac 240
Db 200 gcttcctctctctgtagtgcacatcttctggttagcttccacaaacagtttcaacaattac 259
QY 241 aattcgggacagggagctcctgaagcgttcaagccgagatgagatcatttgcctctgtg 300
Db 260 aattcgggacagggagctcctgaagcgttcaagccgagatgagatcatttgcctctgtg 319
QY 301 ttggttacctcttatttgggaagtgttctctcagcagatcttccacgggtgcactacc 360
Db 320 ttggttacctcttatttgggaagtgttctctcagcagatcttccacgggtgcactacc 379
QY 361 atttcagctatacgaacacactcgaactcttcgcagcaggtgttcttgaatgcaccttgcgg 420
Db 380 atttcagctatacgaacacactcgaactcttcgcagcaggtgttcttgaatgcaccttgcgg 439
QY 421 caagaggtctggcatttcgacaagcagagcaatggagcaatcgccactcargtcaactacc 480
Db 440 caagaggtctggcatttcgacaagcagagcaatggagcaatcgccactcargtcaactacc 499
QY 481 aatggcaacgctatacaaacaggtattgcgagaaatttggcttaccgtgcaggcaactt 540
Db 500 aatggcaacgctatacaaacaggtattgcgagaaatttggcttaccgtgcaggcaactt 559
QY 541 tcaatgtctcttctgcatttggctgcttggcgtctcagtggaagctagctttaaacc 600
Db 560 tcaatgtctcttctgcatttggctgcttggcgtctcagtggaagctagctttaaacc 619
QY 601 accatgtccgtatccctgcatttctctggttcacccggcatctgcataattgagtc 660
Db 620 accatgtccgtatccctgcatttctggttcacccggcatctgcataattgagtc 679
QY 661 gctcagagccagagatcaccaagatctactacgcgcgcgtctctcgcagaagaagtc 720
Db 680 gctcagagccagagatcaccaagatctactacgcgcgcgtctctcgcagaagaagtc 739
QY 721 ttatcatcatcgcgacagtccttctctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
Db 740 ttatcatcatcgcgacagtccttctctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 799
QY 781 gatgtcttttgcagcaagcacacacaaaggaagggaagaaatcgccaaat 831
Db 800 gatgtcttttgcagcaagcacacacaaaggaagggaagaaatcgccaaat 850

RESULT 13
US-09-351-823-9
; Sequence 9, Application US/09351823
; GENERAL INFORMATION:
; APPLICANT: DuVick, Jon
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification
; FILE REFERENCE: 5718-56
; CURRENT APPLICATION NUMBER: US/09/351,823
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 05:29:34 ; Search time 705 Seconds  
(without alignments)  
9234.808 Million cell updates/sec

Title: US-09-882-694A-10  
Perfect score: 3792  
Sequence: 1 atggcagatgaatcgagaa.....aatcttttagaccaagcatga 3792

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3792	100.0	3792	22 AAD02695	Exophiala spinifer
2	780	20.6	3999	22 AAD02694	Exophiala spinifer
3	21	0.6	4002	21 AAX85823	cDNA encoding mult
C 4	20	0.5	359	21 AAA69756	Human ovarian carc
C 5	20	0.5	490	22 AAL16237	Human breast cance
C 6	20	0.5	702	22 AAL25080	Human breast cance
7	20	0.5	6346	18 AAV74548	Staphylococcus aur
8	19	0.5	236	22 AAK75234	Human immune/haema
9	19	0.5	247	22 AAK59415	Human immune/haema

10	19	0.5	354	22 AAH67255	C glutamicum codin
11	19	0.5	477	22 AAF71942	Corynebacterium gl
12	19	0.5	504	22 AAK75237	Human immune/haema
C 13	19	0.5	829	21 AAC34294	Arabidopsis thalia
C 14	19	0.5	1224	22 AAH53092	S. epidermidis ope
15	19	0.5	1449	23 AAS56128	Salmonella typhi D
16	19	0.5	1494	22 AAK82911	Human immune/haema
17	19	0.5	1595	22 AAH55049	S. epidermidis gen
C 18	19	0.5	2390	22 AAD08046	Human extracellular
C 19	19	0.5	3296	20 AAX52267	Protein PRO332 CDN
C 20	19	0.5	3296	21 AAC58590	Human PRO332 prote
C 21	19	0.5	3296	22 AAS21428	Human cDNA sequenc
C 22	19	0.5	3296	22 AAF72425	Human PRO332 cDNA
C 23	19	0.5	3738	22 AAH54614	S. epidermidis gen
C 24	19	0.5	3768	21 AAL15953	Human protein clon
C 25	19	0.5	3800	23 ABL17372	Drosophila melanog
C 26	19	0.5	4046	22 AAH54855	S. epidermidis gen
C 27	19	0.5	12123	22 AAS39855	Genomic sequence #
C 28	19	0.5	12123	22 AAK90226	Human digestive sy
C 29	19	0.5	22788	23 ABL20208	Drosophila melanog
C 30	19	0.5	24999	22 AAK65622	Human immune/haema
C 31	19	0.5	49431	23 ABL13742	Drosophila melanog
C 32	19	0.5	349980	22 AAH68531	C glutamicum codin
C 33	18	0.5	39	16 AAO81380	Forward primer for
C 34	18	0.5	39	18 AAT49185	Murine IgG heavy c
C 35	18	0.5	39	21 AAC84070	Murine Ig heavy ch
C 36	18	0.5	252	14 AAK39882	Expressed Sequence
C 37	18	0.5	252	14 AAK59094	Human brain Expres
C 38	18	0.5	400	22 AAF64561	Novel human polynu
C 39	18	0.5	411	22 AAF66406	Novel human polynu
C 40	18	0.5	432	22 AAH99092	Human EST-derived
41	18	0.5	444	21 AAA43171	Xenopus secreted e
C 42	18	0.5	449	23 AAS68856	DNA encoding novel
C 43	18	0.5	449	23 AAS73984	DNA encoding novel
C 44	18	0.5	455	19 AAV65225	DNA encoding a S.
45	18	0.5	613	16 AAT05457	Human cystic fibro

## ALIGNMENTS

RESULT 1  
AAD02695  
ID AAD02695 standard; DNA; 3792 BP.  
XX AAD02695;  
AC AAD02695;  
XX 02-MAY-2001 (first entry)  
DT Exophiala spinifera p-glycoprotein coding sequence.  
DE Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
KW aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;  
KW detoxification; mycotoxin; animal feed; human feed; silage;  
KW transgenic plant; transgenic animal; microbial spray; ds.  
OS Exophiala spinifera.  
XX Key Location/Qualifiers  
FH 1..3792  
CDS /\*tag= a  
FT /product= "E. spinifera p-glycoprotein"  
FT  
XX WO200105980-A1.  
PN 25-JAN-2001.  
XX 14-JUL-1999; 99WO-US15824.  
PF 14-JUL-1999; 99WO-US15824.  
PR 14-JUL-1999; 99WO-US15824.  
XX (PION-) PIONEER HI-BRED INT INC.  
PA (CURA-) CURAGEN CORP.









```
FT CDS 1..4002
FT /*tag= a
FT /note= "no termination codon given"
XX
XX PN WO9932608-A1.
XX
XX PD 01-JUL-1999.
XX
XX PF 22-DEC-1998; 98WO-US27300.
XX
XX PR 23-DEC-1997; 97US-0996545.
XX
XX PA (ELIL ) LILLY & CO ELI.
XX PA (UYWA-) UNIV WAGENINGEN AGRIC.
XX
XX PI Andrade AC, De Waard MA, Peery RB, Skatrud PL;
XX
XX DR WPI; 1999-418917/35.
XX DR P-PSDB; AAY02630.
XX
XX PT New multiple drug resistance gene atrD
XX
XX PS Claim 2; Page 21-27; 35pp; English.
XX
XX CC The present sequence encodes a multiple drug resistance protein,
XX CC designated atrD, which is isolated from Aspergillus nidulans. The
XX CC atrD polynucleotide is used to produce nucleic acid molecules
XX CC which are useful in diagnostic assays for the detection of atrD
XX CC polynucleotides. The atrD polypeptide may be used to identify agents
XX CC which inhibit multiple drug resistance. These agents may be useful
XX CC for the treatment of fungal diseases. New antifungal compounds may
XX CC also be identified using the gene disruption or gene replacement
XX CC strains of Aspergillus nidulans.
XX
XX SQ Sequence 4002 BP; 974 A; 1034 C; 1007 G; 987 T; 0 other;

Query Match 0.6%; Score 21; DB 20; Length 4002;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3533 gaagcgacgtcagccctcgac 3573
Dy 3763 gaagcgacgtcagccctcgac 3783

RESULT 4
AAA69756/c
ID AAA69756 standard; cDNA; 359 BP.
XX
XX AC AAA69756;
XX
XX DT 07-NOV-2000 (first entry)
XX
XX DE Human ovarian carcinoma antigen polynucleotide SEQ ID NO:66.
XX
XX KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
XX KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200036107-A2.
XX
XX PD 22-JUN-2000.
XX
XX PF 17-DEC-1999; 99WO-US30270.
XX
XX PR 17-DEC-1998; 98US-0215681.
XX PR 17-DEC-1998; 98US-0216003.
XX PR 23-JUN-1999; 99US-0338933.
XX PR 24-SEP-1999; 99US-0404879.
XX
XX PA (CORI-) CORIXA CORP.
```

```
XX
XX PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
XX DR WPI; 2000-431589/37.
XX
XX PT Immunogenic portion of an ovarian carcinoma protein and the nucleic
XX PT acid encoding it, useful for the diagnosis, prevention and treatment of
XX PT cancer, preferably ovarian cancer -
XX
XX PS Claim 1; Fig 1; 299pp; English.
XX
XX CC The present invention describes an isolated polypeptide comprising an
XX CC immunogenic portion of an ovarian carcinoma protein (or its variants).
XX CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
XX CC cytostatic activity and can be used in gene therapy and vaccines.
XX CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
XX CC are useful for the prevention, diagnosis and treatment of cancer,
XX CC preferably ovarian cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557
XX CC represent human ovarian carcinoma polynucleotides and proteins used in
XX CC the exemplification of the present invention.
XX
XX SQ Sequence 359 BP; 125 A; 69 C; 82 G; 82 T; 1 other;

Query Match 0.5%; Score 20; DB 21; Length 359;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2209 ttatatattctgatgttctt 2228
Dy 287 TTTTATATCTGATGTTCTT 268

RESULT 5
AAL16237/c
ID AAL16237 standard; cDNA; 490 BP.
XX
XX AC AAL16237;
XX
XX DT 07-DEC-2001 (first entry)
XX
XX DE Human breast cancer expressed polynucleotide 8694.
XX
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200151628-A2.
XX
XX PD 19-JUL-2001.
XX
XX PF 10-JAN-2001; 2001WO-US00798.
XX
XX PR 14-JAN-2000; 2000US-0176077.
XX PR 14-MAR-2000; 2000US-0189167.
XX PR 24-MAR-2000; 2000US-0192099.
XX PR 29-MAR-2000; 2000US-0193480.
XX PR 15-MAY-2000; 2000US-0205230.
XX PR 09-JUN-2000; 2000US-0211315.
XX PR 25-JUL-2000; 2000US-0220534.
XX
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX DR WPI; 2001-451856/48.
XX
XX PR New peptide useful as a marker for the diagnosis of breast cancer -
XX
XX PS Claim 1; Page 1573; 3695pp; English.
XX
XX CC The invention relates to human breast cancer expressed polynucleotides
XX CC (AAL07544-AAL26789) and methods of assessing whether a patient is
```

CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and encoded  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity.

XX Sequence 490 BP; 154 A; 107 C; 106 G; 121 T; 2 other;

Query Match 0.5%; Score 20; DB 22; Length 490;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2209 ttattatctgattctt 2228

Db 187 TTTTATATCTGATGTTCTT 168

RESULT 6

AAAL25080/c  
 ID AAL25080 standard; cDNA; 702 BP.

XX AAL25080;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 17537.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US00798.

XX 14-JAN-2000; 2000US-0176077.

XX 14-MAR-2000; 2000US-0189167.

XX 24-MAR-2000; 2000US-0192099.

XX 29-MAR-2000; 2000US-0193480.

XX 15-MAY-2000; 2000US-0205230.

XX 09-JUN-2000; 2000US-0211315.

XX 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer -  
 XX Claim 1; Page 3241-3242; 3695pp; English.  
 CC The invention relates to human breast cancer expressed polynucleotides  
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity.

XX Sequence 702 BP; 225 A; 149 C; 158 G; 170 T; 0 other;

Query Match

Best Local Similarity 100.0%; Pred. No. 18;

0.5%; Score 20; DB 22; Length 702;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2209 ttattatctgattctt 2228

Db 165 TTTTATATCTGATGTTCTT 146

RESULT 7

AAV74548  
 ID AAV74548 standard; DNA; 6346 BP.

XX AAV74548;

XX 16-MAR-1999 (first entry)

XX Staphylococcus aureus contig SEQ ID #237.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;  
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
 KW skin infection; surgical wound infection; scalded skin syndrome;  
 KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

XX Key Location/Qualifiers

FT misc\_feature 1081..1140

FT /\*tag= a

FT /note= "these bases represent a line of missing text in  
 the sequence listing in the specification. They  
 are included to maintain the nucleotide numbering  
 given in the specification for this DNA sequence."

FT misc\_feature 2881..2940

FT /\*tag= b

FT /note= "these bases represent a line of missing text in  
 the sequence listing in the specification. They  
 are included to maintain the nucleotide numbering  
 given in the specification for this DNA sequence."

FT misc\_feature 4681..4740

FT /\*tag= c

FT /note= "these bases represent a line of missing text in  
 the sequence listing in the specification. They  
 are included to maintain the nucleotide numbering  
 given in the specification for this DNA sequence."

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-0100117.

XX 05-JAN-1996; 96US-0009861.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;  
 PI Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
 PT stored on computer readable medium and used in the production of  
 PT anti-S.aureus vaccines

XX Claim 1; Page 1017-1020; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
 CC of the invention. The DNA sequences are recorded on a computer readable  
 CC medium, preferably selected from a floppy or hard disk, random access  
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
 CC the S.aureus DNA sequences allows putative functions to be assigned so  
 CC that protein-encoding or regulatory regions of commercial, therapeutic or  
 CC industrial importance can be obtained. Specifically, sequences which are  
 CC likely to encode antigens have been identified and these polypeptides can

CC be used in a vaccine composition against *S. aureus* infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the *S. aureus* DNA sequences contained on the  
CC computer readable medium.

XX Sequence 6346 BP; 2170 A; 914 C; 1229 G; 1849 T; 184 other;

Query Match 0.5%; Score 20; DB 18; Length 6346;  
Best Local Similarity 100.08; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 638 goactgcatagcaattgat 657  
|||||  
Db 4152 goactgcatagcaattgat 4171

## RESULT 8

AAK75234  
ID AAK75234 standard; DNA; 236 BP.

XX AAK75234;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30046.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytosolic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214896.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0217496.

XX 26-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220563.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0241826.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.



```
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244817.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX P-PSDB; AAM86634.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Claim 1; SEQ ID NO 4475; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 247 BP; 65 A; 67 C; 56 G; 57 T; 2 other;

Query Match 0.5%; Score 19; DB 22; Length 247;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 792 gcagcaagcacaccacaaga 810
   |||||
Db 129 gcagcaagcacaccacaaga 147

RESULT 10
AAH67255
ID AAH67255 standard; DNA; 354 BP.
XX
XX AAH67255;
XX
XX 26-SEP-2001 (first entry)
XX
XX C glutamicum coding sequence fragment SEQ ID NO: 2290.
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
OS
```

XX PN EP1108790-A2.  
XX PD 20-JUN-2001.  
XX PF 18-DEC-2000; 2000EP-0127688.  
XX PR 16-DEC-1999; 99JP-0377484.  
XX PR 07-APR-2000; 2000JP-0159162.  
XX PR 03-AUG-2000; 2000JP-0280988.  
XX PA (KYOW ) KYOWA HAKKO KOCYO KK.  
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX PR WPI; 2001-376931/40.  
XX DR P-PSDB; AAG92036.  
XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying  
XX PT mutation point of a gene, measuring expression of a gene, analysing  
XX PT expression profile or pattern of a gene and identifying homologous gene  
XX PS Claim 8; SEQ ID NO: 2290; 246pp + Sequence Listing; English.  
XX CC The present invention provides a number of nucleotide and protein  
XX CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
XX CC are useful for identifying the mutation point of a gene derived from a  
XX CC mutant of coryneform bacterium, measuring expression amount and  
XX CC analysing the expression profile or expression pattern of a gene derived  
XX CC from Coryneform bacterium, and identifying a homologue of a gene derived  
XX CC from coryneform bacterium. Coryneform bacteria are useful for producing  
XX CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
XX CC particularly L-lysine. The present sequence is a nucleic acid described  
XX CC in the exemplification of the invention.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from the  
XX CC European Patent Office.  
XX SQ Sequence 354 BP; 83 A; 103 C; 98 G; 70 T; 0 other;  
  
Query Match 0.5%; Score 19; DB 22; Length 354;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2329 caacgagtactggtatcaag 2347  
Db 195 caacgagtactggtatcaag 213  
  
RESULT 11  
AAAF71942  
ID AAF71942 standard; DNA; 477 BP.  
AC AAF71942;  
XX 30-APR-2001 (first entry)  
XX DT  
XX DE  
XX KW Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:379.  
XX KW fine chemical production; metabolic pathway protein; MP protein;  
XX KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
XX KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
XX KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.  
XX OS Corynebacterium glutamicum.  
XX PN WO200100843-A2.  
XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-IB00923.  
XX PR 25-JUN-1999; 99US-0141031.  
XX PR 01-JUL-1999; 99DE-1030476.  
XX PR 02-JUL-1999; 99US-0142101.  
XX PR 08-JUL-1999; 99DE-1031415.  
XX PR 08-JUL-1999; 99DE-1031418.  
XX PR 08-JUL-1999; 99DE-1031419.  
XX PR 08-JUL-1999; 99DE-1031420.  
XX PR 08-JUL-1999; 99DE-1031424.  
XX PR 08-JUL-1999; 99DE-1031428.  
XX PR 08-JUL-1999; 99DE-1031434.  
XX PR 08-JUL-1999; 99DE-1031435.  
XX PR 08-JUL-1999; 99DE-1031443.  
XX PR 08-JUL-1999; 99DE-1031453.  
XX PR 08-JUL-1999; 99DE-1031457.  
XX PR 08-JUL-1999; 99DE-1031478.  
XX PR 08-JUL-1999; 99DE-1031510.  
XX PR 08-JUL-1999; 99DE-1031541.  
XX PR 08-JUL-1999; 99DE-1031573.  
XX PR 08-JUL-1999; 99DE-1031592.  
XX PR 08-JUL-1999; 99DE-1031632.  
XX PR 08-JUL-1999; 99DE-1031634.  
XX PR 08-JUL-1999; 99DE-1031636.  
XX PR 09-JUL-1999; 99DE-1032125.  
XX PR 09-JUL-1999; 99DE-1032126.  
XX PR 09-JUL-1999; 99DE-1032130.  
XX PR 09-JUL-1999; 99DE-1032186.  
XX PR 09-JUL-1999; 99DE-1032206.  
XX PR 09-JUL-1999; 99DE-1032227.  
XX PR 09-JUL-1999; 99DE-1032228.  
XX PR 09-JUL-1999; 99DE-1032229.  
XX PR 09-JUL-1999; 99DE-1032230.  
XX PR 14-JUL-1999; 99DE-1032922.  
XX PR 14-JUL-1999; 99DE-1032926.  
XX PR 14-JUL-1999; 99DE-1032928.  
XX PR 14-JUL-1999; 99DE-1033004.  
XX PR 14-JUL-1999; 99DE-1033005.  
XX PR 14-JUL-1999; 99DE-1033006.  
XX PR 12-AUG-1999; 99US-0148613.  
XX PR 27-AUG-1999; 99DE-1040764.  
XX PR 27-AUG-1999; 99DE-1040765.  
XX PR 27-AUG-1999; 99DE-1040766.  
XX PR 27-AUG-1999; 99DE-1040832.  
XX PR 31-AUG-1999; 99DE-1041378.  
XX PR 31-AUG-1999; 99DE-1041379.  
XX PR 31-AUG-1999; 99DE-1041380.  
XX PR 31-AUG-1999; 99DE-1041384.  
XX PR 31-AUG-1999; 99DE-1041396.  
XX PR 03-SEP-1999; 99DE-1042076.  
XX PR 03-SEP-1999; 99DE-1042077.  
XX PR 03-SEP-1999; 99DE-1042079.  
XX PR 03-SEP-1999; 99DE-1042086.  
XX PR 03-SEP-1999; 99DE-1042087.  
XX PR 03-SEP-1999; 99DE-1042088.  
XX PR 03-SEP-1999; 99DE-1042095.  
XX PR 03-SEP-1999; 99DE-1042124.  
XX PR 03-SEP-1999; 99DE-1042129.  
XX PR 09-MAR-2000; 2000US-0187970.  
XX PA (BADI ) BASF AG.  
XX PI Pompeius M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
XX PR WPI; 2001-137957/14.  
XX DR P-PSDB; AAB79823.  
XX PT Nucleic acids from Corynebacterium glutamicum encoding metabolic  
XX PT pathway proteins, useful for producing fine chemicals in  
XX PT microorganisms, including organic acids, nonproteinogenic amino acids,  
XX PT and purine and pyrimidine bases -



```
XX Claim 3; Page 723-724; 1737pp; English.
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX Sequence 477 BP; 123 A; 132 C; 130 G; 92 T; 0 other;
SQ

Query Match          0.5%; Score 19; DB 22; Length 477;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2329 caacgagtactggatcaag 2347
Db 295 caacgagtactggatcaag 313
|||||

RESULT 12
AAK75237
ID AAK75237 standard; DNA; 504 BP.
XX AC AAK75237;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30049.
XX KW Human; immune; haematopoietic; Immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-APR-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
```

PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and  
XX metastasis.  
XX Disclosure; SEQ ID NO 30049; 3071pp + Sequence Listing; English.  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and  
XX treatment of diseases associated with inappropriate (I) expression. For  
XX example, they may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of (I) by expressing inactive proteins or to  
XX supplement the patient's own production of (I). Additionally, (I)  
XX polynucleotides may be used to produce the secreted (I), by inserting  
XX the nucleic acids into a host cell and culturing the cell to express the  
XX protein. (I) proteins and polynucleotides may be used to prevent,  
XX diagnose and treat immune/haematopoietic-related diseases, especially  
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
XX to AAK87694 represent human immune/haematopoietic antigen genomic  
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
XX represent sequences used in the exemplification of the present invention.  
XX Sequence 504 BP; 139 A; 129 C; 114 G; 122 T; 0 other;  
Query Match 0.5%; Score 19; DB 22; Length 504;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 792 gcagcaagcaccacacagaa 810

Db 192 gcagcaagcaccacacagaa 210  
|||||  
RESULT 13  
AAC34294/c  
ID AAC34294 standard; DNA; 829 BP.  
XX  
AC AAC34294;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 6135.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP103340S-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 23-MAR-1999; 99US-0123548.  
PR 25-MAR-1999; 99US-0125788.  
PR 29-MAR-1999; 99US-0126264.  
PR 01-APR-1999; 99US-0126785.  
PR 06-APR-1999; 99US-0127462.  
PR 16-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 19-APR-1999; 99US-0129845.  
PR 21-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140895.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 0.5%; Score 19; DB 21; Length 829;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 tggactctcaatgcatcg 154  
|||||  
Db 52 TGGACTCTCAATGTCATCG 34

RESULT 14  
AAH53092/c  
ID AAH53092 standard; DNA; 1224 BP.

```
XX AAH53092;
AC
XX
XX 03-SEP-2001 (first entry)
DT
XX
XX S. epidermidis open reading frame nucleotide sequence SEQ ID NO:1577.
DE
XX
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW
KW vaccination; endocarditis; ds.
XX
XX Staphylococcus epidermidis.
OS
XX
XX WO200134809-A2.
PN
XX
XX 17-MAY-2001.
PD
XX
XX 09-NOV-2000; 2000WO-US30782.
PF
XX
XX 09-NOV-1999; 99US-0164258.
PR
XX
XX (GLAX ) GLAXO GROUP LTD.
PA
XX
XX Kimmerly WJ;
PI
XX
XX WPI; 2001-316495/33.
DR
XX
XX P-PSDB; AAG82242.
DR
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT
XX useful for vaccinating against infections, e.g. endocarditis -
PT
XX
XX Claim 8; Page 443; 2189pp; English.
XX
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC
CC (I) and (II) can have antibacterial activity and therefore can be used
CC
CC in vaccination. The nucleic acids (I) may be used to produce the
CC
CC S. epidermidis polypeptides (II) via the production of vectors
CC
CC containing them which are used to produce hosts cells which express the
CC
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC
CC The polypeptides may also be used to assay for other inhibitors of their
CC
CC activity and therefore identify compounds that may be used for the
CC
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC
CC polynucleotide sequences from the present invention. AAH55091 to
CC
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC
CC in the exemplification of the present invention.
CC
CC N.B. The present invention specifically claims all the polynucleotide
CC
CC sequences given in the sequence listing of the present specification,
CC
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
XX Sequence 1224 BP; 384 A; 199 C; 219 G; 422 T; 0 other;
SQ
```

```
Query Match 0.5%; Score 19; DB 22; Length 1224;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1303 atattacttgatgggtgg 1321
|||||
Db 857 ATATTACTTGATGGTGG 839
```

```
RESULT 15
AA56128
ID AA56128 standard; DNA; 1449 BP.
XX
AC AA56128;
XX
XX 13-FEB-2002 (first entry)
DT
XX
```

```
DE
XX Salmonella typhi DNA for cellular proliferation protein #161.
XX
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW
KW antibiotic; antibacterial; drug design.
XX
XX Salmonella typhi.
OS
XX
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR
XX
XX 23-MAY-2000; 2000US-206848P.
PR
XX
XX 26-MAY-2000; 2000US-207737P.
PR
XX
XX 23-OCT-2000; 2000US-242578P.
PR
XX
XX 27-NOV-2000; 2000US-253625P.
PR
XX
XX 22-DEC-2000; 2000US-257931P.
PR
XX
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR
XX
XX P-PSDB; AAU38269.
DR
XX
XX New polynucleotides for the identification and development of
PT
XX antibiotics, comprise sequences of antisense nucleic acids -
PT
XX
XX Claim 27; Seq ID No 9765; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC
XX prokaryotic cellular proliferation, their use in identifying the
CC
XX genes, their use in the discovery of novel antibiotics, the essential
CC
XX genes themselves and the encoded proteins. The prokaryotes used are
CC
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC
XX invention is also useful for the identification of potential new targets
CC
XX for antibiotic development. The antisense nucleic acids can also be used
CC
XX to identify proteins used in proliferation, to express these proteins,
CC
XX and to obtain antibodies capable of binding to the expressed proteins.
CC
XX The proteins can be used to screen compounds in rational drug discovery
CC
XX programmes. The antisense nucleic acid sequence is also useful to screen
CC
XX for homologous nucleic acids which are required for cell proliferation in
CC
XX a wide variety of organisms. The present sequence encodes an
CC
XX essential prokaryotic cellular proliferation protein.
CC
XX Note: The sequence data for this patent did not form part
CC
XX of the printed specification, but was obtained in electronic
CC
XX format directly from WIPO at
CC
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1449 BP; 372 A; 350 C; 402 G; 309 T; 16 other;
SQ
```

```
Query Match 0.5%; Score 19; DB 23; Length 1449;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2620 tgcagaaagtgcctggcctt 2638
|||||
Db 518 tgcagaaagtgcctggcctt 536
```

Search completed: April 27, 2002, 05:30:14  
Job time: 13897 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2002, 05:44:33 ; Search time 7820.58 Seconds  
(without alignments)  
10146.744 Million cell updates/sec

Title: US-09-882-694A-10

Perfect score: 3792  
Sequence: 1 atgcagatgaatcgaggaa.....aatcttagaccaagcatga 3792

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl :

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	3792	100.0	3792	6	AX076851
2	780	20.6	3999	6	AX076850
3	23	0.6	128	9	HC0146S19
4	22	0.6	10104	1	AE005604
5	22	0.6	1414	1	AE000450
6	22	0.6	14526	1	ECUNC
7	22	0.6	136254	1	ECOUN82
8	22	0.6	280900	1	AP002566
9	21	0.6	824	10	MMPOLE22
10	21	0.6	1362	10	BC002275
11	21	0.6	2348	5	AF193800
12	21	0.6	2879	5	AB030443
13	21	0.6	4002	6	ARI49776
14	21	0.6	4002	6	ARI49777
15	21	0.6	4569	8	AF173826
16	21	0.6	6133	8	AF071411
17	21	0.6	7119	10	AF123502
18	21	0.6	10297	1	AE006358
19	21	0.6	37566	1	MSG8971CS
20	21	0.6	73463	9	AL161785
21	21	0.6	73995	2	AC109140
22	21	0.6	82905	2	AC106950
23	21	0.6	83946	9	HSA123M24
24	21	0.6	89479	8	AC006932
25	21	0.6	97912	9	HSJ753D10
26	21	0.6	102533	9	AL390115
27	21	0.6	110000	2	AC087331_1
28	21	0.6	154357	2	AC098689
29	21	0.6	155699	9	AL360176
30	21	0.6	165292	2	AL590113
31	21	0.6	167475	2	AC106083
32	21	0.5	1204	9	BC017424
33	20	0.5	1590	1	AF338160
34	20	0.5	1770	3	TVPGP1
35	20	0.5	3353	10	MUSGCE
36	20	0.5	11695	1	AE005117
37	20	0.5	35961	2	AC108249
38	20	0.5	49336	1	AC025950
39	20	0.5	49822	2	AC011038
40	20	0.5	50152	1	AC025948
41	20	0.5	57650	2	AC095828
42	20	0.5	66619	2	AC100648
43	20	0.5	76042	9	AL357084
44	20	0.5	78941	9	AC005591
45	20	0.5	78973	8	AB012242

## ALIGNMENTS

RESULT	1	AX076851	Sequence 10 from Patent WO0105980.	3792 bp	DNA	linear	PAT 22-FEB-2001
LOCUS	AX076851	Sequence 10 from Patent WO0105980.					
DEFINITION	AX076851	Sequence 10 from Patent WO0105980.					
ACCESSION	AX076851	Sequence 10 from Patent WO0105980.					
VERSION	AX076851.1	GI:13121523					
KEYWORDS	Exophiala spinifera.						
SOURCE	Exophiala spinifera						
ORGANISM	Exophiala spinifera						
REFERENCE	1 (bases 1 to 3792)						
AUTHORS	Duwick,J.P., Maddox,J., Gilliam,J., Folkerts,O. and Crasta,O.R.						
TITLE	Compositions and methods for fumonisins detoxification						
JOURNAL	Patent: WO 0105980-A 10 25-JAN-2001;						
FEATURES	Pioneer HI-Bred International Inc. (US) ; Curagen Corporation (US)						
LOCATION/Qualifiers	1..3792						
ORGANISM	/organism="Exophiala spinifera"						
DB_xref	/db_xref="taxon:91928"						
note	/note="p-glycoprotein, coding sequence"						
BASE COUNT	939 a	983 c	928 g	940 t			2 others

ORIGIN									
Query Match 100.0%; Score 3792; DB 6; Length 3792; Best Local Similarity 100.0%; Pred. No. 0; Matches 3792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	atg	c	c	a	a	a	a	g
Db	1	atg	c	c	a	a	a	a	g
Qy	61	c	c	t	c	c	c	c	a
Db	61	c	c	t	c	c	c	c	a
Qy	121	g	c	c	a	a	a	a	t
Db	121	g	c	c	a	a	a	a	t
Qy	181	g	c	t	c	c	c	c	a
Db	181	g	c	t	c	c	c	c	a
Qy	241	a	a	t	c	c	c	c	a
Db	241	a	a	t	c	c	c	c	a
Qy	301	t	c	c	a	a	a	a	t
Db	301	t	c	c	a	a	a	a	t
Qy	361	a	t	c	c	a	a	a	t
Db	361	a	t	c	c	a	a	a	t
Qy	421	c	a	a	a	a	a	a	t
Db	421	c	a	a	a	a	a	a	t
Qy	481	a	a	t	c	c	c	c	a
Db	481	a	a	t	c	c	c	c	a
Qy	541	t	c	c	a	a	a	a	t
Db	541	t	c	c	a	a	a	a	t
Qy	601	a	c	c	a	a	a	a	t
Db	601	a	c	c	a	a	a	a	t
Qy	661	g	c	c	a	a	a	a	t
Db	661	g	c	c	a	a	a	a	t
Qy	721	t	a	t	c	c	a	a	a
Db	721	t	a	t	c	c	a	a	a
Qy	781	g	a	t	c	c	a	a	a
Db	781	g	a	t	c	c	a	a	a
Qy	841	t	t	c	c	a	a	a	a
Db	841	t	t	c	c	a	a	a	a
Qy	901	t	t	c	c	a	a	a	a
Db	901	t	t	c	c	a	a	a	a
Qy	961	t	c	c	a	a	a	a	a
Db	961	t	c	c	a	a	a	a	a
Qy	1021	t	c	c	a	a	a	a	a
Db	1021	t	c	c	a	a	a	a	a
Qy	1081	t	c	c	a	a	a	a	a
Db	1081	t	c	c	a	a	a	a	a
Qy	1141	t	c	c	a	a	a	a	a
Db	1141	t	c	c	a	a	a	a	a
Qy	1201	t	c	c	a	a	a	a	a
Db	1201	t	c	c	a	a	a	a	a
Qy	1261	t	c	c	a	a	a	a	a
Db	1261	t	c	c	a	a	a	a	a
Qy	1321	t	c	c	a	a	a	a	a
Db	1321	t	c	c	a	a	a	a	a
Qy	1381	t	c	c	a	a	a	a	a
Db	1381	t	c	c	a	a	a	a	a
Qy	1441	t	c	c	a	a	a	a	a
Db	1441	t	c	c	a	a	a	a	a
Qy	1501	t	c	c	a	a	a	a	a
Db	1501	t	c	c	a	a	a	a	a
Qy	1561	t	c	c	a	a	a	a	a
Db	1561	t	c	c	a	a	a	a	a
Qy	1621	t	c	c	a	a	a	a	a
Db	1621	t	c	c	a	a	a	a	a
Qy	1681	t	c	c	a	a	a	a	a
Db	1681	t	c	c	a	a	a	a	a
Qy	1741	t	c	c	a	a	a	a	a
Db	1741	t	c	c	a	a	a	a	a
Qy	1801	t	c	c	a	a	a	a	a
Db	1801	t	c	c	a	a	a	a	a
Qy	1861	t	c	c	a	a	a	a	a
Db	1861	t	c	c	a	a	a	a	a
Qy	1921	t	c	c	a	a	a	a	a
Db	1921	t	c	c	a	a	a	a	a
Qy	1981	t	c	c	a	a	a	a	a
Db	1981	t	c	c	a	a	a	a	a
Qy	2041	t	c	c	a	a	a	a	a
Db	2041	t	c	c	a	a	a	a	a
Qy	2100	t	c	c	a	a	a	a	a
Db	2100	t	c	c	a	a	a	a	a

QY 2101 ataacgttctgatatgcgagccacacattccaggcaaacccctttgttttcgagattg 2160  
|||||  
Db 2101 ATAACGGTCTTGATATCGCGGCCACATTTCCAGGACAAGCCCTTTTGTTCGAGATTG 2160  
QY 2161 ctcaactctctcagttgagtggaatgcgagcagagaaacggcagacattttatctg 2220  
Db 2161 CTCACGTCTTCGAGTTGAGTGGTCATCGCGSCACAGAACGGCGACACTTTTATATCTG 2220  
QY 2221 atgttctcttgcgtggccttaggaatctagtaggattttcacgattgctgagacatgc 2280  
|||||  
Db 2221 ATGTTCTTTGCGTCTAGGAAATCTAGTAGGATTTTACGATTGGCTGGACATGC 2280  
QY 2281 aacgttattcacaagttgtccaccatcgctatcaacgcgaatgtttccaacagtagtctg 2340  
Db 2281 AACGTTATTTTACAAAGTTGTCAACCATCGCTATCAAGCCGCAATTTTCCAAACGAGTACTG 2340  
QY 2341 gatcaagacatcgaactctcgacatcccgagcaaatcttctgctctcacatcgcaa 2400  
|||||  
Db 2341 GATCAAGACATCGAATCTCGATCCCGAGCAAAATTTCTGGTCTCTCACATCGCAA 2400  
QY 2401 ctgtcagctctaccacgcagttgcaagagttgatatcagcaaatctttctcatttatct 2460  
Db 2401 CTGTCAGCTCTACCCACGAGTTGCAAGAGTTGATATCAGCAAAATTTTCTCAATTATATC 2460  
QY 2461 gttgtcgttaacatcgctcagcagtgctctaccactagcctatggatggaactggg 2520  
Db 2461 GTTGTGCGTCAACATCGTCTCGAGCAGTGTCTTACCCTAGCTATGGATGGAACTGGG 2520  
QY 2521 cctggatgttgggtgcaactccacccctgcttttggctggcttaacctgaattcg 2580  
Db 2521 CCTGGTGTGTGTTGGTGACATCTCCACCCCTGCTTTTGGCTGGCTACCTCAGAAATCG 2580  
QY 2581 tctagagcgaagtagaagccggaactcggcaactcttgcagaaagtgtcgggttgc 2640  
Db 2581 TCTAGAGACGAAGTAGAAGCCGGAACCTCGGCAAACTTTTGCAGAAAGTGTCTGGGCTGC 2640  
QY 2641 aagcgaagcagttaccgcgactccgagccgtctcactcttgaactctcgaagccatctct 2700  
Db 2641 AAGCGAAGCAGTTACCGGATCCGGACCGTCTCATCTTTGACTCTCGAAGCCATGTTCT 2700  
QY 2701 ccaacagtaactcagacatgttgagcaaggtcttgcgaagatcatccaaagcttttggtt 2760  
Db 2701 CCAACAGTACTCGGACATGTTGAGCAAGTCTTGTCTAAGATCATCCAAAGCTTTTGGTTT 2760  
QY 2761 gacagattttgttttcttctcactctcagatcgatgagtttttggctattgacctggga 2820  
Db 2761 GACGATGTTTGGTTTCTACCTGCTCAGTGTGATGGATTTTGGCTATTTGGCTTGGGA 2820  
QY 2821 tttgtattgcagtcgataattggcttcaggtgagtcagcacacaactcaattttatatca 2880  
Db 2821 TTTTGTATTGTCAGTCGATAATTGGCTTCAGTGTGAGTACGACACAACCTCAATTTATATCA 2880  
QY 2881 tcttcgtggggttttcttgcggtccagcagcagcagcagcagcagcagcagcagcagcag 2940  
Db 2881 TCTTCTGGGGGGTGTGTTTGGCGGTCCAAAGCAGCAGCCAGTATTGGCTTACTCCAGC 2940  
QY 2941 agttttaccagagctcgtcgtgcgaactatctctggtcgtgcgacattgaaagcgcg 3000  
Db 2941 AGTTTACCAGGCTCGTGGCTGCGCAACTATATCTCTGGCTGCGGACATTTGAGCCG 3000  
QY 3001 accatccgcgaacgagggagaaacagaaaaagccagtggggtggatgcccctgtcgac 3060  
Db 3001 ACCATCCCGAAACGAGGAGAAACAGAAAAAGGCGCCAGTGGTGGATGCCCTGTGAC 3060  
QY 3061 ctcaagagcattgaatlcaggtatcgtcaacgtgattcgactcgagttctcccggggtt 3120  
Db 3061 CTCGAGGACATTAATTCAGTATCTGTCACAGTGAATTCGGCTCGAGTCTCCCGGGGT 3120  
QY 3121 tccatgacaatcagagcaggagaaattgttagcttatgtggcgctcttgctggtggcaag 3180  
Db 3121 TCCATGACAATCGAGCCAGGACAAATTTGTAGCTTATGTGGCGCTTCTGCGCTGTGCAAG 3180

QY 3181 tcaacgttgcagctttgtcggaaagattcttacgacccgacctcgggcgaaatttcattt 3240  
|||||  
Db 3181 TCAACGTTGATCGCTTGTTCGGAAGATTCTACGACCGGACCTCGGGCGGAATTTCAATTT 3240  
QY 3241 gcaacgcgaataattgcgaaatgtcgcgcgtttgtacgcggccatatgtctttggctc 3300  
Db 3241 GCACACGAGAATATTGCAAAATGTCCCGGCTTGTACCGCGGCCATATGCTCTTTGGTC 3300  
QY 3301 caacaggaaccacacactttaccaaggtccgtctcgagagaatgtgacgttgccctcgaa 3360  
Db 3301 CAACAGGAACCCACACTTTACCAAGGCTCCGTTCCGAGAAATGTGAGTGGCCCTCGAA 3360  
QY 3361 gccgaattatcagaagagctttgtcaaggagcgttcccgcaagggcccaatgctttggatt 3420  
Db 3361 GCCGAATATCAGAAGAGCTTTGTCAAGGAGCGCTTCCGCAAGGCAATGCTTTTGGATT 3420  
QY 3421 ttgtcatctctttaccagaagctttgaaagcgttcgagccttcgagctcaacgagatcgattc 3480  
Db 3421 TTGTCTATCTCTTTACCAGAAGCTTTGAACGCTTTCGGCTCAACGAGGATCGAGTTC 3480  
QY 3481 tccgcgggcaacgacgagctgcgcatcgccatcgcaagagcattgattccaaatccaaagctg 3540  
Db 3481 TCCGCGGGCAACGACGAGCGGATCGCCATCGCAAGAGCATTGATTCGAAATCCCAAAGCTG 3540  
QY 3541 ttgctacttgacgaagacgagctcagccctcgacacgaagaatcggaacgtctgtttcaagct 3600  
Db 3541 TTGTCTATCTCTTTACCAGAAGCTTCAGCCCTCGACACGCAATCGGAACGCTTGTGGTTCAAGCT 3600  
QY 3601 gccctcgatgagcgtccacagcagccgcaacgacaatagcagtgccgacgactttccact 3660  
Db 3601 GCCCTCGATGAGGATCCACGAGCGGACGACAAATAGCAGTGGCGCACGACTTTTCCACT 3660  
QY 3661 attcggaattgtatgttt 3720  
Db 3661 ATTCTGGAATGTTGATGTTTATTTTGTGTTTGTTCGCAACGGAGAAATCGCGCAACGGGCACT 3720  
QY 3721 cagcggaactcacaacgactgagaggaagatatlacagatgtgtttggcacatcttta 3780  
Db 3721 CACGCGGAACCTACACGACTGAGAGGAAGATATTACGAGATGTTTGGCACACATCTTTTA 3780  
QY 3781 gaccaagcatga 3792  
Db 3781 GACCAAGCATGA 3792

## RESULT 2

AX076850 LOCUS AX076850 3999 bp DNA linear PAT 22-FEB-2001  
DEFINITION Sequence 9 from Patent WO0105980.  
ACCESSION AX076850  
VERSION AX076850.1 GI:13121522  
KEYWORDS  
SOURCE Exophiala spinifera.  
ORGANISM Exophiala spinifera  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothyriomycetes;  
Chaetothiales; Herpotrichiellaceae; mitosporic  
Herpotrichiellaceae; Exophiala.  
REFERENCE 1 (bases 1 to 3999)  
AUTHORS Duvick J. P., Maddox J., Gilliam J., Folkerts O. and Crasta, O.R.  
TITLE Compositions and methods for fumonisin detoxification  
JOURNAL Patent: WO 0105980-A 9 25-JAN-2001;  
Pioneer Hi-Bred International Inc. (US); Curagen Corporation (US)  
FEATURES  
source  
/organism="Exophiala spinifera"  
/db\_xref="taxon:91928"  
/note="P-glycoprotein, with introns"  
BASE COUNT 1007 a 1008 c 992 g 987 t 5 others  
ORIGIN

Query Match 20.6%; Score 780; DB 6; Length 3999;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches 830; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggcagatgaatcgagaaactcgacaaacaaagatggcagtgagtcgtccacac 60  
 Db 20 atggcagatgaatcgagaaactcgacaaacaaagatggcagtgagtcgtccacac 79  
 QY 61 cctcccccagaaagaaacgaagcagtgatttcagactatctacgaatcttcagatat 120  
 Db 80 cctcccccagaaagaaacgaagcagtgatttcagactatctacgaatcttcagatat 139  
 QY 121 gccgacaatacagactgacitcctaataatcgcgcctcatctcgccatcgatccggg 180  
 Db 140 gccgacaatacagactgacitcctaataatcgcgcctcatctcgccatcgatccggg 199  
 QY 181 gcttcccttctctgatgcgatcatcttcggtagcttcacaaacaaagttcaacaattac 240  
 Db 200 gcttcccttctctgatgcgatcatcttcggtagcttcacaaacaaagttcaacaattac 259  
 QY 241 aattcggcgacggagtcctgaagcgttcgaagccgatgtggatcatcttcgctctgg 300  
 Db 260 aattcggcgacggagtcctgaagcgttcgaagccgatgtggatcatcttcgctctgg 319  
 QY 301 ttcgctaccctcttattggaaattgtctcctcagctacgtttccacggctgcattacc 360  
 Db 320 ttcgctaccctcttattggaaattgtctcctcagctacgtttccacggctgcattacc 379  
 QY 361 attcagctatcacaaacactcgaactcttcgacgagtgcttccttgatgcactctgcg 420  
 Db 380 attcagctatcacaaacactcgaactcttcgacgagtgcttccttgatgcactctgcg 439  
 QY 421 caagaggtctggtcatttcgacaaacagacgaatggcaatcgccactcargtcactacc 480  
 Db 440 caagaggtctggtcatttcgacaaacagacgaatggcaatcgccactcargtcactacc 499  
 QY 481 aatggcaaccgtatatacaaacaggtattgcgagaaatgtcttaccggtgcaggaactt 540  
 Db 500 aatggcaaccgtatatacaaacaggtattgcgagaaatgtcttaccggtgcaggaactt 559  
 QY 541 tcaatgtctttctgcatcttggctgcttggcgttcagtggaagctagctttaatc 600  
 Db 560 tcaatgtctttctgcatcttggctgcttggcgttcagtggaagctagctttaatc 619  
 QY 601 accatgtcgtatccttgccatttcttgatcagcgcctgtctcctcgagaaagtc 660  
 Db 620 accatgtcgtatccttgccatttcttgatcagcgcctgtctcctcgagaaagtc 679  
 QY 661 gctcaggagggcagatcacagagatctactcagcgcctgtctcctcgagaaagtc 720  
 Db 680 gctcaggagggcagatcacagagatctactcagcgcctgtctcctcgagaaagtc 739  
 QY 721 ttatcatcatcagcagtcctgcttcttctacgccagaaagaaatggtcgaaaaatat 780  
 Db 740 ttatcatcatcagcagtcctgcttcttctacgccagaaagaaatggtcgaaaaatat 799  
 QY 781 gatgtcttttgcagcagcacacaaagaaagaaagaaatcgccaaat 831  
 Db 800 gatgtcttttgcagcagcacacaaagaaagaaagaaatcgccaaat 850

RESULT 3  
 LOCUS HC04A6S19/C  
 DEFINITION Human type IV collagen (COL4A6) gene, exon 18.  
 ACCESSION U46977  
 VERSION U46977.1 GI:1674411  
 KEYWORDS .  
 SEGMENT 19 of 46  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 128)  
 AUTHORS Zhang, X., Zhou, J., Reiders, S.T. and Tryggvason, K.

Structure of the human type IV collagen COL4A6 gene, which is mutated in Alport syndrome-associated leiomyomatosis  
 Genomics 33 (3), 473-479 (1996)  
 96299642  
 2 (bases 1 to 128)  
 Zhang, X.  
 Direct Submission  
 Submitted (24-JAN-1996) Xu Zhang, MBB, Karolinska, Stockholm, 171-77, Sweden  
 Location/Qualifiers  
 1. 128  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /map="Xq22"  
 /cell\_line="1635; contains multiple X-chromosomes"  
 11. 118  
 /gene="COL4A6"  
 /number=18  
 33 a 33 c 34 g 28 t  
 BASE COUNT  
 ORIGIN  
 exon  
 Query Match 0.6%; Score 23; DB 9; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 0.92;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 666 ggagggcaggatcacaggatct 688  
 Db 43 GGAGGCCAGGATCACCAGGATCT 21  
 RESULT 4  
 LOCUS AE005604  
 DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 223 of 290.  
 ACCESSION AE005604  
 VERSION AE005604.1 GI:12518574  
 KEYWORDS  
 SOURCE Escherichia coli O157:H7 EDL933.  
 ORGANISM Escherichia coli O157:H7 EDL933.  
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.  
 REFERENCE 1 (bases 1 to 10104)  
 AUTHORS Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.  
 Genome sequence of enterohemorrhagic Escherichia coli O157:H7  
 Nature 409 (6819), 529-533 (2001)  
 21074935  
 MEDLINE 11206551  
 REFERENCE 2 (bases 1 to 10104)  
 AUTHORS Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.  
 Direct Submission  
 Submitted (23-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA  
 Location/Qualifiers  
 1. 10104  
 /organism="Escherichia coli O157:H7 EDL933"  
 /strain="EDL933"  
 /serotype="O157:H7"  
 /db\_xref="taxon:155864"  
 /note="enterohemorrhagic complement(74. 1114)  
 gene





```

/db_xref="GI:12518581"
/translation="MKPMIVGALALTSVFMAGHLQAAADGVVHFRGEIIDSCTEVTPE
TKDQVDLGVKVRFAESGDDVAAAPTAFSIDLTCCPEFKSAAIFREDGNEADAHGNGNL
AIGTPLDNSNDAAAGISPSDSNGDYTGAGVSAKGVAIRLYNRADNTQVKLYENSAS
TPISNGASMKFMARIATETITDPTGTANADSQFTVEYIK"
complement(7777. .7866)
/gene="25226"
complement(7777. .7866)
/gene="25226"
/function="orf; Unknown function"
/note="No significant matches"
/codon_start=1
/transl_table=11
/protein_id="AAG89931.1"
/db_xref="GI:12518582"

```

Query Match 0.6%; Score 22; DB 1; Length 10104;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 2452 atttatatcggttgctggccaac 2473
|||||
Db 10009 ATTATATCGTGTGCGTCAAC 10030

```

## RESULT 5

```

AE000450 11414 bp DNA linear BCT 01-DEC-2000
LOCUS Escherichia coli K12 MG1655 section 340 of 400 of the complete
DEFINITION genome.
ACCESSION AE000450 U00096
VERSION AE000450.1 GI:1790166
KEYWORDS
SOURCE Escherichia coli K12.
ORGANISM Escherichia coli K12.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

```

```

1 (bases 1 to 11414)
Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
Mau,B. and Shao,Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (5331), 1453-1474 (1997)
97426617
9278503

```

2 (bases 1 to 11414)  
 Blattner,F.R.  
 Direct Submission  
 Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,  
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.  
 Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:  
 608-263-7459

3 (bases 1 to 11414)  
 Blattner,F.R.  
 Direct Submission  
 Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,  
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.  
 Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:  
 608-263-7459

4 (bases 1 to 11414)  
 Blattner,F.R.  
 Direct Submission  
 Submitted (13-OCT-1998) Laboratory of Genetics, University of  
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

This sequence was determined by the E. coli Genome Project at the  
 University of Wisconsin-Madison (Frederick R. Blattner, director).  
 Supported by NIH grants HG00301 and HG01428 (from the Human Genome  
 Project and NCHGR). The entire sequence was independently  
 determined from E. coli K12 strain MG1655. Predicted open reading  
 frames were determined using Genemark software, kindly supplied by  
 Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,  
 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

Submitted (13-OCT-1998) Laboratory of Genetics, University of  
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA  
 This sequence was determined by the E. coli Genome Project at the  
 University of Wisconsin-Madison (Frederick R. Blattner, director).  
 Supported by NIH grants HG00301 and HG01428 (from the Human Genome  
 Project and NCHGR). The entire sequence was independently  
 determined from E. coli K12 strain MG1655. Predicted open reading  
 frames were determined using Genemark software, kindly supplied by  
 Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,  
 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that

have been correlated with genetic loci are being annotated with CG  
 Site Nos., unique ID nos. for the genes in the E. coli Genetic  
 Stock Center (CGSC) database at Yale University, kindly supplied by  
 Mary Berlyn. A public version of the database is accessible  
 (<http://cgsc.biology.yale.edu>). Annotation of the genome is an  
 ongoing task whose goal is to make the genome sequence more useful  
 by correlating it with other data. Comments to the authors are  
 appreciated. Updated information will be available at the E. coli  
 Genome Project's World Wide Web site  
 (<http://www.genetics.wisc.edu>). \*\*\* The E. coli K12 sequence and  
 its annotations are periodically updated; this is version #54. No  
 sequence changes. Annotation updates: updated gene identifications  
 and products; all new functional assignments courtesy of Monica  
 Riley; added promoters, protein binding sites, and repeated  
 sequences described in reference 1. The unique numeric identifiers  
 beginning with a lowercase 'b' assigned to each gene (protein- or  
 RNA-encoding) are now designated as gene synonyms instead of  
 labels. This should allow them to be searched for in Entrez as gene  
 names.

## FEATURES

Location/Qualifiers	source
1. .11414 /organism="Escherichia coli K12" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333" complement(29. .61) /notes="factor Sigma70; promoter phos; documented +1 at 3909196"	promoter
complement(45. .62) /notes="central position to phoS promoter: -15" /bound_moiety="PhoB documented site" complement(66. .83) /notes="central position to predicted promoter: -36; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(67. .84) /notes="central position to phoS promoter: -37" /bound_moiety="PhoB documented site" complement(67. .84) /notes="central position to predicted promoter: -37; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(77. .94) /notes="central position to predicted promoter: -47; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(292. .2121) /genes="gimS" /notes="b3729" complement(292. .2121) /genes="gimS" /function="enzyme; Central intermediary metabolism: Amino sugars" /notes="f609; 99 pct identical amino acid sequence and equal length to GLMS_ECOLI SW: P17169; CG Site No. 706" /codon_start=1 /transl_table=11 /product="L-glutamine:D-fructose-6-phosphate aminotransferase" /protein_id="AAC76752.1" /db_xref="GI:1790167"	protein_bind
complement(67. .84) /notes="central position to phoS promoter: -37" /bound_moiety="PhoB documented site" complement(67. .84) /notes="central position to predicted promoter: -37; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(77. .94) /notes="central position to predicted promoter: -47; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(292. .2121) /genes="gimS" /notes="b3729" complement(292. .2121) /genes="gimS" /function="enzyme; Central intermediary metabolism: Amino sugars" /notes="f609; 99 pct identical amino acid sequence and equal length to GLMS_ECOLI SW: P17169; CG Site No. 706" /codon_start=1 /transl_table=11 /product="L-glutamine:D-fructose-6-phosphate aminotransferase" /protein_id="AAC76752.1" /db_xref="GI:1790167"	protein_bind
complement(67. .84) /notes="central position to phoS promoter: -37" /bound_moiety="PhoB documented site" complement(67. .84) /notes="central position to predicted promoter: -37; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(77. .94) /notes="central position to predicted promoter: -47; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(292. .2121) /genes="gimS" /notes="b3729" complement(292. .2121) /genes="gimS" /function="enzyme; Central intermediary metabolism: Amino sugars" /notes="f609; 99 pct identical amino acid sequence and equal length to GLMS_ECOLI SW: P17169; CG Site No. 706" /codon_start=1 /transl_table=11 /product="L-glutamine:D-fructose-6-phosphate aminotransferase" /protein_id="AAC76752.1" /db_xref="GI:1790167"	protein_bind
complement(67. .84) /notes="central position to phoS promoter: -37" /bound_moiety="PhoB documented site" complement(67. .84) /notes="central position to predicted promoter: -37; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(77. .94) /notes="central position to predicted promoter: -47; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(292. .2121) /genes="gimS" /notes="b3729" complement(292. .2121) /genes="gimS" /function="enzyme; Central intermediary metabolism: Amino sugars" /notes="f609; 99 pct identical amino acid sequence and equal length to GLMS_ECOLI SW: P17169; CG Site No. 706" /codon_start=1 /transl_table=11 /product="L-glutamine:D-fructose-6-phosphate aminotransferase" /protein_id="AAC76752.1" /db_xref="GI:1790167"	gene
complement(67. .84) /notes="central position to phoS promoter: -37" /bound_moiety="PhoB documented site" complement(67. .84) /notes="central position to predicted promoter: -37; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(77. .94) /notes="central position to predicted promoter: -47; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(292. .2121) /genes="gimS" /notes="b3729" complement(292. .2121) /genes="gimS" /function="enzyme; Central intermediary metabolism: Amino sugars" /notes="f609; 99 pct identical amino acid sequence and equal length to GLMS_ECOLI SW: P17169; CG Site No. 706" /codon_start=1 /transl_table=11 /product="L-glutamine:D-fructose-6-phosphate aminotransferase" /protein_id="AAC76752.1" /db_xref="GI:1790167"	CDS
complement(67. .84) /notes="central position to phoS promoter: -37" /bound_moiety="PhoB documented site" complement(67. .84) /notes="central position to predicted promoter: -37; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(77. .94) /notes="central position to predicted promoter: -47; genetic evidence for the site" /bound_moiety="PhoB predicted site" complement(292. .2121) /genes="gimS" /notes="b3729" complement(292. .2121) /genes="gimS" /function="enzyme; Central intermediary metabolism: Amino sugars" /notes="f609; 99 pct identical amino acid sequence and equal length to GLMS_ECOLI SW: P17169; CG Site No. 706" /codon_start=1 /transl_table=11 /product="L-glutamine:D-fructose-6-phosphate aminotransferase" /protein_id="AAC76752.1" /db_xref="GI:1790167"	promoter

```

/note="factor Sigma70; predicted +1 start at 3911406"
complement(2283..3653)
/gene="glnu"
/note="b3730"
complement(2283..3653)
/gene="glnu"
/function="enzyme; Central intermediary metabolism:
Sugar nucleotide biosynthesis, conversions"
/note="f456; ??? pct identical to GLMU_ECOLI SW:
P17114; similar to Bacillus subtilis tms; correction to
previous sequence"
/codon_start=1
/transl_table=11
/product="N-acetyl glucosamine-1-phosphate
uridylyltransferase"
/protein_id="AAC76753.1"
/db_xref="GI:1790168"
/translation="MLNNAMSVILAAKGTRMYSDLPKVLHTLACKAMVQHVDAAN
EUGAHVHLVYGGDLKQALKDDNLNVLQAEQIGTGHAMQCAFPFADDEDILML
YGDVPLISVETLQURDARPOGGIGLLTVKLDOPGTGYRITRENGKVTGIVEHKDAD
EQRIQIEITGLIANGADMKRLAKLTNNAGOEYITDIITAIAYQEGREIVAVHPQ
RISEVQGVNRLQLSLRLERYVSEQAELLAGVMLRDPARFDLRTLPHGRDVEIDT
NVLIEGNTVLGRVKITGCVTKNSVIGDDCETSPYVVEDANLAAACTTGPFAURLP
GAELLEAGVGNFVENKKARLKGSKAGHLTVLGDAEIGDNVNIAGATTTCNYDGANK
FTIIGDDVFVGSOTLVAPTVVGKATTAAGTTVTRNNGENALASRVPQTKESWR
RPVKKK"
complement(2463..2479)
/gene="glnu"
/note="central position to predicted promoter: -224.5"
/bound_moiety="Pdhr predicted site"
complement(3711..3733)
/note="No predicted promoter"
/bound_moiety="NagC documented site"
complement(3864..3886)
/note="No predicted promoter"
/bound_moiety="NagC documented site"
complement(4006..4425)
/gene="atpC"
/note="b3731"
complement(4006..4425)
/gene="atpC"
/EC_number="3.6.1.34"
/function="enzyme; ATP-proton motive force
interconversion"
/note="f139; 100 pct identical to ATPC_ECOLI SW: P00832;
alternate gene names uncC, papG; CG Site No. 31"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
epsilon-subunit"
/protein_id="AAC76754.1"
/db_xref="GI:1790169"
/translation="MAMTVHLDVWSAEQOMFSLGVKIQVTGSEGLIYPGHPALLT
AIKPGMIRIVKOHGHEEFTYLSGILEVOPGNVTVLADTAIRGQDLDEARAMEARKKA
EHISSHGDDVYQASASLAKIAQLRIELTKKAM"
complement(4446..5828)
/gene="atpD"
/note="b3732"
complement(4446..5828)
/gene="atpD"
/EC_number="3.6.1.34"
/function="enzyme; ATP-proton motive force
interconversion"
/note="f460; 100 pct identical to APPB_ECOLI SW: P00824;
alternate gene names uncD, papB; CG Site No. 30"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
beta-subunit"
/protein_id="AAC76755.1"
/db_xref="GI:1790170"
/translation="MATGRIVQVIGAVOVFFQDAPRVYDALEVQNGNERLIVLEVQ
QQLGGGIVTRTIAMGSSDGLRRGLDVRDLEHPTEVPVPGKATLGRIMNVLPFVDMKEI

```

```

GEERWAIHRAAPSYEELSOLFLETGKVIDLMCPFAKGGKVLGFGAGVGKTVNM
MELINIAIEHSGYSVFAGVGRTRGNDFFYHMTDSSNVIDKSVLYGCOMSPGNRL
RVALTGLTNAERFRDEGRDVLFFVDNIYRTIAGTEVSALLGRMPSAVQYQPTLAEM
GVQKERITSTKTSITSVQAVYVPADLDLPPSPATFAHLDATVVLISROIASLGIYPA
VDPLDSTRLQDPLVVGQEHYDTARGVQSIQRYQELKDIITAILGDELSEEDKLVA
EAKTQRLSPQPFVAVAEVFTGSPKGVYSLKDTIRGFKIMEGEYDHLPPQAFYMGSI
EFAVEKAKKL"
complement(4519..4548)
promoter
Query Match 0.6%; Score 22; DB 1; Length 11414;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2452 atttatatcggtcggtcaac 2473
|||||
Db 2136 ATTATATCGTTCGCGTCAAC 2157
|||||

RESULT 6
ECUNC/c 14526 bp DNA linear BCT 29-APR-1998
LOCUS E. coli origin of replication oric and genes gid, unc, ECOURF-1 and
DEFINITION glms.
ACCESSION X01631
VERSION X01631.1 GI:43256
KEYWORDS Inverted repeat; origin of replication; transposable element.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
REFERENCE 1 (bases 1 to 14526)
AUTHORS Walker,J.E., Gay,N.J., Saraste,M. and Eberle,A.N.
TITLE DNA sequence around the Escherichia coli unc operon. Completion of
the sequence of a 17 kilobase segment containing asnA, oric, unc,
glms and phos
JOURNAL The Biochemical Journal. 224 (3), 799-815 (1984)
MEDLINE 85121806
PUBMED 6395859
REFERENCE 2 (bases 1 to 14526)
AUTHORS Messer,W., Hartmann-Kuehlein,H., Langer,U., Mahlow,E., Roth,A.,
Schaper,S., Urmonet,B. and Woelker,B.
TITLE The complex for replication initiation of escherichia coli
JOURNAL Chromosoma 102, 1-6 (1992)
FEATURES
source 1. 14526
/organism="Escherichia coli"
/db_xref="taxon:562"
<1..28
/note="16K protein"
/codon_start=2
/transl_table=11
/protein_id="CAA25772.1"
/db_xref="GI:809693"
/db_xref="SWISS-PROT:P03817"
/translation="GSWNLLK"
65..296
/note="min. region of oric"
343..348
/promoter -35 region"
367..372
/promoter -10 region"
395..398
/note="pot. Shine Dalgarno region"
407..2293
/note="gidA protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25773.1"
/db_xref="GI:43257"
/db_xref="SWISS-PROT:P17112"
/translation="MFYPDPFDVILIGKGAGTGAAMAAARMGQQTLLTLHTIDTLGQ
MSCNPAIGGIGKHLVKVEDLGLGMLKAIDQAGIOFRILNASKGPVATRAQADRV
LYRQAVRTALENPQNLMIQQAAVEDLIVENDRVVGAVTQMKLFRAKAVLVTGTFLD

```

```

GKIHLGDNYSGRAGDPPPSIPLSRRLREPLRVGRKLTGTPPRIDARTIDFVSVAQOQ
HGDNMPVFSMGNASQHPQOVCYITHTNEKTHDVIRSNLDRSPMVAGVTEGVGPY
CPSIDKVRPADRNHOHIFLEPEGLTSNETYPNGISTSLPFDVQMOIVRSMQMGENA
KIVPGVAYEYDFDPDLKPTLESKEFLOGLFFAGQINGTFTGYEANAQGLLAGLNA
RLSADKEGNAFARSQAYLGVLDLCLTGKTEPYRMTISRAEYREMAUQDLRLTE
IGREGLVDDDERWARFNEKENIERERQRLKSTWVTPSAEAAAEVNAHLTAPLSREAS
GEDLLRPMTEYKLTUTTPFAPALTDEQAQAEQVIEQVKEGYIARQQDETEKLRNEN
TLLPATLDYRQVSGLSNEVIATKLNDRKHPASTGQASRISGVTTPAAISILLVWLKQOGL
RRSA"
repeat_unit      2305..2330
/notes="pot. hairpin structure"
RBS              2346..2349 Shine Dalgarno region"
CDS              2357..2980
/notes="gldB protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25774.1"
/db_xref="GI:581247"
/translacion="MLNKLSSLLKLDAGISLTHOKNOLIAVYNMLKHNKAYNLTSVR
DPNMLVRHILDSIVAPYQGERFIDVGTGPGIPGISIVRPEAHFTLLDLSLGRV
FLRQVQHEKLENIEPVQSRVEFPSEPPDFGVISAFASLNDVSWCHLPGEOGR
FYALKGQMPDEITALLPEEQVESVWKLQVPALDGERHLVVIKANKI"
3474..3479
/notes="pot. -35 region"
/notes="pot. -3505"
/notes="pot. -10 region"
3512
/notes="pot. transcription start site (unc)"
3585..3977
/notes="uncI protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25775.1"
/db_xref="GI:581248"
/db_xref="SWISS-PROT:P03808"
/translacion="MKNVSVLSVRNVARKLLVLQVLVVIASGLFLSKDPFWGVS
ISGGLAVFLNVLPMFIAMWRHQHTPAKGRVAVTFAFGEAFKVLAMLVLLVVALAVL
AVFLPLVTWVLVVLVQQLAPAVINNG"
3979..3981
RBS              3986..4801 Shine Dalgarno region"
CDS              3986..4801
/notes="uncB protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25776.1"
/db_xref="GI:43260"
/db_xref="SWISS-PROT:P00855"
/translacion="MASENMTPODYIGHLNNLQDLRTESLVDPPNPATFWTINID
SMFYSVVLGLFLVFLRSVAKKATSGVPKFOFAIELYIGFVNGSVKDMYHGKSLIA
PLATIFVFWFLMLMDLLPIDLPLTAETHVIGLPLRVPSADVNVLTSNALGVFIL
ILFYSIRMKGGIGFTKELTLOPFNHAFIPVNLILEGVLSLSPKPSLGLRFGNMYAG
ELIPILTAGLLPMMWSQILNVPWAIFHILITLQAFIFMWLTIVYLSMASEEH"
4838..4841
RBS              4848..5087 Shine Dalgarno region"
CDS              4848..5087
/notes="uncE protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25777.1"
/db_xref="GI:43261"
/db_xref="SWISS-PROT:P00844"
/translacion="MENLMDLLYMAAAMVMGLAAIGAAIGIGLGGKLEGAARQPD
LIPILRQFTFVIMGLVDAIPMIAGVLGLYVFA"
5145..5149
/notes="pot. Shine Dalgarno region"
RBS              5149..5619
/notes="uncF protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25778.1"
/db_xref="GI:581249"

```

```

/db_xref="SWISS-PROT:P00859"
/translacion="MNLNATILGQAIAFVLFLVLCMKYKVPPLMAAIEKROKEIADGL
ASARAEIKDLDAKASATDOLKKAKAQAQVIEQANRRSOILDEAKAAEQERTKIV
AQQAELAEARKRAREELRQVALAVAGAERKILERSVDEANSDIVDKLVAEL"
5619..5622 Shine Dalgarno region"
/notes="pot. Shine Dalgarno region"
5625..5628
/notes="pot. Shine Dalgarno region"
5634..6167
/notes="unch protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25779.1"
/db_xref="GI:43263"
/db_xref="SWISS-PROT:P00831"
/translacion="MSEFIVTAPYAKAADFVAFHOSVSRWQDMLAFAAEVTKEQM
AELLSGALAPETLAESTIAVCGEOLDENGNLIRVMAENGRNLNLPVLEQFIHLRAV
SEATAEVDVISAALSQQLAKISAAMEKRLSRVKLNCKIDKSVMAGVIIRAGDMVI
DGSVRGLERLADVLQS"
6175..6178
/notes="pot. Shine Dalgarno region"
RBS              6180..7721
CDS              6180..7721
/notes="uncA protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25780.1"
/db_xref="GI:43264"
/db_xref="SWISS-PROT:P00822"
/translacion="MQLNSTEISELISIKORIAQFNVVSEAHNEGTIVSVDGVIHGL
ADCQGMISLPGNYAIALNLSERDSVAVMGYPADLAEGMKYKCTGRILEVPGVR
LLGRVNTLGPIDGKGLDIDGFSAVEATAPYIERQSDQVPQYGVKVSIMPTG
RGRELIGDROTKTALADAILNQRDSGKICLYVAIGQKASTISNVVKLEHGAL
ANTIVVATASEAALQYLAARMPVALMGEYFRDRGEDALIIYDOLSKQAAVAYRQTSLL
LRPEGREAPGDVFLHLSRLLEHAAVNAEYVEATFKGEVKGTGSLTALPILETOA
GDVSAFPTNVISITDQIFLETNLPNAGIRPAPNPGISVSRVGGAGTAKMKLSGC
IRYALAGTRELAFSQFASDLDNATRAQDLQDQKAYTELKOKQAPMSVAAQSULFA
AERYADLVLSKIGTSFEAALLAYVDRHAPLMOEINOTGYNDEIEGKLAGILDSEK
ATQSGM"
7726..7753
repeat_unit     7726..7764 hairpin structure"
RBS              7761..7764
CDS              7772..8635 Shine Dalgarno region"
/notes="uncG protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25781.1"
/db_xref="GI:43265"
/db_xref="SWISS-PROT:P00837"
/translacion="MACAKEIRSKIASVONTOKITKAMEMVAASMKRSODRMAASRP
YAEIMRVIGHLAHNGLEYKHPYLEDROVKRQVYLVSTDRGLGGLNINLFKLLAE
MKWTDRKGVCDLHAGICKGVSFNSVGGNNVAVOVTGMDNPSLSSELIGPVKVMLOAY
DEGRDLKLIYVSNKFIINTMSQVPTISQLPLPDDDDDLKHKSWDYLTERDPAKLLDT
LLRRYVESQYQGVYENLASEQAARMYAKAANDTNGGSLIKELQLVYNKARQASITOE
LTEIVSGAAAV"
8652..8655
RBS              8662..10044 Shine Dalgarno region"
CDS              8662..10044
/notes="uncD protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25782.1"
/db_xref="GI:899257"
/db_xref="SWISS-PROT:P00824"
/translacion="MATGKIVQVIGAVGVDFEPODAPRVYVDELVONGNERLVLEVO
QOLGGIVRTIAMCSDGLRGLDKLEHPVEPVGKATLGRIMNVLSEPDVMKGEI
GEERWAIHRAAPSYELSNLSQELLETKIVIDLMCFPAKGGKVLGGGAGVGTVM
MELIRNIAIEHSGYSFVAGVGTRENDYFHEMTSDVSDKSLVYGMNPEPPNRL
RVALTGLTMAEKFRDEGRDVLFDVNIYRTYTLAGTESALIGRMPSPVAGYOPTLAEM
GVQERTSTKTGTSITQAVYVPADLTDPSPTATFAHLDATVLLSRQIASLGIYPA
VPLDSTSRQDPLVQGEHYDARGVQSLQRYQELKDIILILGMDLSSEDLKVA
RARKIQRFLSOPFEVAFVFTGSPGKYVSLKDTIRGFKGIMEGEVDHLPEQAFYMGVSI
EEAVEKAKKL"

```

```

RBS      10057..10059
Query Match      0.6%; Score 22; DB 1; Length 14526;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2452 atttatactgttcggtcaac 2473
|||||
Db 12354 ATTATATCGTTCGCGTCAAC 12333

RESULT 7
ECOW82      136254 bp      DNA      linear      BCT 07-FEB-1995
LOCUS      E. coli; the region from 81.5 to 84.5 minutes.
DEFINITION      L10328
ACCESSION      L10328
VERSION      L10328.1 GI:290484
KEYWORDS
SOURCE
ORGANISM      Escherichia coli K12 strain MG1655; lambda clones EC14-52,
EC17-187, EC27-890, EC21-99, EC22-175, EC27-236, EC17-30, EC17-137,
EC19-71, EC15-15, EC17-159; subclones in M13mpl9 or Janus.
Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 136254)
Burland,V., Plunkett,G. III, Daniels,D.L. and Blattner,F.R.
DNA sequence and analysis of 136 kilobases of the Escherichia coli
genome: organizational symmetry around the origin of replication
Genomics 16 (3), 551-561 (1993)
93315143
This sequence was determined as part of the E. coli Genome Project
(Frederick R. Blattner, director) at the University of
Wisconsin-Madison. Supported by award HG00301 from the NIH Human
Genome Project. A preliminary report was presented at the Hilcon
Head meeting/Genome Sequencing and Analysis iv' September 26 -
30,1992. The entire sequence was independently determined from E.
coli MG1655; overlaps and conflicts with other sequence
determinations are annotated. The end of this entry overlaps the
start of the entry ECOW85 (M87049) by the six bases of an EcoRI
site.
Data kindly submitted in computer readable form by: Guy Plunkett
III
Laboratory of Genetics
University of Wisconsin
445 Henry Mall
Madison WI 53706
USA
Phone: 608-262-2534
Email: ecoligenetics.wisc.edu
Fax: 608-263-7459
[1] Authors request hold until publication.

FEATURES
     source      1..136254
                  Location/Qualifiers
                    /organism="Escherichia coli"
                    /db_xref="taxon:562"
                    complement(<1..669)
                    /note="X06036; ECOFFG(1..>669)"
                    1..11607
                    /note="EC14-52 (lambda clone)"
                    complement(<1..544)
                    /gene="fpg"
                    /note="alternate gene name muth"
                    /codon_start=1
                    /transl_table=11
                    /product="formamidopyrimidine-DNA glycosylase"
                    /protein_id="AAA61988.1"
                    /db_xref="GI:643609"
                    /translation="MPPEVETSRRGIEPHLVGATILHAVVNGRLRWVPSEIYRL
SQDVLVSORAKYLILLEPGEWIIHLGMSGSLRILPELPEKIDHDVLYMSKV
LRYDPRRGAWLWTELEGNHNLTHLGPEPLSDDFNGEYLHOKCAKKKTAIKPLWMD
NKLIVGVGNIVASELSFAAGI"
                    complement(1..544)
                    /gene="fpg"

     misc_feature      complement(526..1290)
                        /note="J01677; ECORPMRG(1..764)"
     promoter           complement(557..586)
                        /note="promoter-like sequence; promoter matrix score of
                        44"
     misc_difference    596..601
                        /note="AAAAA in X06036; AAAAAA in J01677 and here"
     misc_difference    611
                        /note="G in J01677; A in X06036 and here"
     misc_difference    615
                        /note="G in J01677; A in X06036 and here"
     misc_difference    618..623
                        /note="TTTTTT in X06036; TTTTTT in J01677 and here"
     gene              complement(642..809)
                        /gene="rpmG (CG Site No. 236)"
     CDS               complement(642..809)
                        /gene="rpmG (CG Site No. 236)"
                        /codon_start=1
                        /transl_table=11
                        /product="50S ribosomal subunit protein L33"
                        /protein_id="AAA61989.1"
                        /db_xref="GI:290486"
                        /translation="MAKGIREKIKLYSSAGTGHFYTTTKNKTKEPKLELKKFDPVVR
                        QHVIYKEAKIK"
                        complement(830..1066)
                        /gene="rpmB (CG Site No. 240)"
                        complement(830..1066)
                        /gene="rpmB (CG Site No. 240)"
                        /codon_start=1
                        /transl_table=11
                        /product="50S ribosomal subunit protein L28"
                        /protein_id="AAA61990.1"
                        /db_xref="GI:290487"
                        /translation="MSRVCQVTGKRPVTGNNRSHALNATKRRFLPNLHSHRWFESEK
                        RFVLRVSAKGMRVIDKRGIDTVLAEALARGEKY"
     misc_difference    902
                        /note="G in J01677; A here"
     misc_difference    1128
                        /note="TC in J01677; C here"
     misc_difference    1153..1155
                        /note="T in J01677; TAT here"
                        complement(1168..1819)
                        /note="X63366; ECORADC(1..651)"
                        complement(1212..1240)
                        /note="promoter-like sequence; promoter matrix score of
                        77"
     gene              complement(1283..1957)
                        /gene="radC (CG Site No. 13913)"
                        complement(1283..1957)
                        /gene="radC (CG Site No. 13913)"
                        /note="differs from radC sequence in X63366; similar to S.
                        aureus putative 25 kb protein"
                        /codon_start=1
                        /transl_table=11
                        /protein_id="AAA61991.1"
                        /db_xref="GI:290488"
                        /translation="MKYKNSQLLMPREKMLKFGISALTDYELLALFLYCTGRGDVL
                        TLAKEMLENFGSLYGLTSEYEQFSGVHGIGYAKPAQKLAELARRIYNVRREES
                        LLSPEMREFLOSQTGEERIEFMVIFLDQSQRVITHRRLFSGLTNHVEVHPREIRE
                        AIKINASALILAHNHPSCABPSKADKLITERIKSKQCFMDLRVLDHIVIGRGYVSE
                        AERGWI"
     misc_difference    1369..1370
                        /note="A in X63366; AA here"
     variation          1436
                        /note="A deleted in radC102 mutation (X63367)"
     misc_structure     1698..1822
                        /note="predicted bend of 76 degrees"
     promoter           1875..1902
                        /note="promoter-like sequence; promoter matrix score of
                        50"
     promoter           complement(1911..1940)
                        /gene="radC (CG Site No. 13913)"
                        /note="promoter-like sequence; promoter matrix score of

```

gene	45"	2051..3343	/gene="dfp (CG Site No. 18430)"
	CDS	2051..3343	/gene="dfp (CG Site No. 18430)"
misc_feature			/codon_start=1
			/transl_table=11
misc_feature			/protein_id="AAA61992.1"
			/db_xref="GI:290489"
gene			/translation="MKARQKYCDKIANFWCHPTGKIIMSLAGKKIVLGVSGGTAAYK
			TPELVRLRGADVRVAMTEAKAFITPLSLQAVSGIPVSDSLDDPAEAAMGHIEL
CDS			GKADVLIPATADTLARVAAGMANDLVSTICLATPAPVAVLPAMQOQMYRAAATQH
			NLEYLAGRLLINGLDSGSGACGDIGPGRXXDPLTIYDMVAHFSPVNDLKLHLMIT
misc_feature			AGTPREDPVRYISNHSKMGFAIAAARAGANTLVSGPVSLSLPTPPVVKRVDM
			TALEMEAAVNASVQQNITFGCAAVADYRAATVAPEKTKKQATGDELTIKWKNPDI
misc_feature			VAGVAALKDRHPYVYVGAETNNVEEYARQKRIRKNDLILICANDVSQPTQGFNSDNN
			LHLPWQGDVPLERKELGQLLDELIVTRYDEKRR"
misc_feature			2982..4591
			/note="X01714; ECOPYREA(1..1609)"
misc_feature			2982..5552
			/note="V01578; ECODUTPYR(1..2568)"
gene			3324..3779
			/gene="dut (CG Site No. 832)"
CDS			3324..3779
			/gene="dut (CG Site No. 832)"
misc_difference			/codon_start=1
			/transl_table=11
repeat_region			/product="deoxyuridine triphosphatase"
			/protein_id="AAA61993.1"
CDS			/db_xref="GI:290490"
			/translation="MKKIDVKILDPYRGKEPLPTATYSSAGLDLACRLNDVAVELAP
misc_difference			GDTLVPTGLAIHIDPDSLAAMLPRSGLGKHGIVLGNLVGLDSDYQGQMLTSVWN
			RGDSFTIQPCERIAQMLFVPUQAERNLVDFDQTRDGGEGFGHSGRQ"
CDS			3844..4482
			/gene="ttk"
misc_difference			3844..4482
			/gene="ttk"
repeat_region			/function="unknown"
			/codon_start=1
CDS			/transl_table=11
			/product="23 kD protein"
misc_difference			/protein_id="AAA61994.1"
			/db_xref="GI:290491"
repeat_region			/translation="MPFGKCLFSGVFCNMAEKQAKRNRREIIQLSLMLESDDGSGQ
			RITPAKAASGVSEALYRHFPSKTFMDFSLIEDLSLITRNLILKDEKDTTARL
CDS			RLIVLLGLGFCERNPGITRIULTGHALMFEODRLQGRINOLFETEAOLROYLREKRM
			EGEGYTTDELLASQLACFEGMLSRFVRSEFKYRPTDDFDRWPLIAAQLQ"
misc_difference			4469..4471
			/gene="ttk"
repeat_region			/note="CC in X01714 and V01578; CAC here"
			4489..4514
gene			/note="REP (repetitive extragenic palindromic) element;
			contains 1 REP sequence"
CDS			complement(4519..5160)
			/gene="pyrE (CG Site No. 327)"
misc_difference			complement(4519..5160)
			/gene="pyrE (CG Site No. 327)"
repeat_region			/codon_start=1
			/transl_table=11
CDS			/product="rotate phosphoribosyltransferase"
			/protein_id="AAA61995.1"
misc_difference			/db_xref="GI:543610"
			/translation="MKPYQROFIEFALSQVLFKEGFTLKSGRKSPYFNFAGLENTGR
CDS			DLAGLGRFYAELVDGSPGKIPATTTAVALAEHDLDDLPYCFNRKEA
			KDHGEGNVLGSAQLQGMVLDVITAGTAIRESMEIIQANGATLAVLISLDROERG
misc_difference			RGEISAIQVEYRDYCNKVISIITLKDLIAYLEEKPEMAHLAAVKAAYREERG"
			4762..4766
misc_difference			0.6%; Score 22; DB 1; Length 136254;
			Query Match
misc_difference			100.0%; Pred. No. 3.5;
			Best Local Similarity
misc_difference			0; Mismatches 0; Indels 0; Gaps 0;
			Matches 22; Conservative
misc_difference			2452 atttatatcgttgctggccaac 2473
			Qy

Db 103074	ATTTATATCGTTGTCGTCAC 103095
RESULT	8
	AP002566
LOCUS	Escherichia coli O157:H7 DNA, complete genome, section 17/20.
	AP002566 BA000007
ACCESSION	AP002566.1 GI:13363930
	Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RMD 0509952)
VERSION	DNA.
	Escherichia coli O157:H7
KEYWORDS	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
	Escherichia.
SOURCE	1 (sites)
	Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
ORGANISM	Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
	Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
REFERENCE	Sasakawa,C. and Shinagawa,H.
	Complete nucleotide sequence of the prophage VT2-Sakai carrying the
AUTHORS	verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
	derived from the Sakai outbreak
TITLE	Genes Genet. Syst. 74 (5), 227-239 (1999)
	20198780
JOURNAL	2 (sites)
	Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M.,
MEDLINE	Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and
	Hayashi,T.
REFERENCE	Comparative analysis of the whole set of rRNA operons between an
	enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
AUTHORS	Escherichia coli K-12 strain MG1655
	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
TITLE	20557356
	3 (sites)
JOURNAL	Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S.,
	Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T.,
MEDLINE	Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and
	Shinagawa,H.
REFERENCE	Complete nucleotide sequence of the prophage VT1-Sakai carrying the
	Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
AUTHORS	O157:H7 strain derived from the Sakai outbreak
	Gene 238 (1-2), 127-139 (2000)
TITLE	20564182
	4 (sites)
JOURNAL	Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K.,
	Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T.,
MEDLINE	Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C.,
	Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and
REFERENCE	Shinagawa,H.
	Complete genome sequence of enterohemorrhagic Escherichia coli
AUTHORS	O157:H7 and genomic comparison with a laboratory strain K-12
	DNA Res. 8 (1), 11-22 (2001)
TITLE	21156231
	5 (bases 1 to 280900)
JOURNAL	Ohnishi,M., Kurokawa,K., Makino,K., Yasunaga,T., Shinagawa,H. and
	Hayashi,T.
MEDLINE	Direct Submission
	Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
REFERENCE	Information Research Center: 3-1, Yamadaoka, Suita, Osaka 565-0871,
	Japan [E-mail:ken@gen-info.osaka-u.ac.jp,
AUTHORS	URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
	Fax:81-6-6879-2047)
COMMENT	genome project.
	Location/Qualifiers
FEATURES	1..280900
	/organism="Escherichia coli O157:H7"
source	/strain="O157:H7"
	/sub_strain="RMD 0509952"
gene	/db_xref="taxon:83334"
	complement(143..616)
gene	/gene="ECs4456"

```

CDS
complement(143..616)
/ gene="ECs4456"
/ note="similar to YIAI_ECOLI YIAI_ECOLI gi|2367245 percent identity 98 in 157 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="hypothetical protein"
/protein_id="BAB3789.1"
/db_xref="GI:13363931"
/transl_table=11
/translation="MNRPIADATKICGCTCEVACAVSHQENQDCAALSPDEFISRI
RVTKHDSWTTAVACHOCEDAPCANVCPDAISREHGHIFVEQTCIGCKSCMLACPGF
AMEVSSKKARAIKDCLCWHRETGPACVACPTKALQCMDEKVKQRHLRQPV"
complement(860..1675)
/ gene="ECs4457"
complement(860..1675)
/ gene="ECs4457"
/ note="probable transcriptional regulator, weakly similar to hypothetical transcriptional regulator Yisr [bacillus subtilis] gi|3123306|spIP40331 percent identity 24 in 276 aa"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="putative transcriptional regulator"
/protein_id="BAB37880.1"
/db_xref="GI:13363932"
/translation="MLELSMTLPKVGNGGLFISRGVGRHSARRLESWEIIFVEKGR
LTIQEEERFLVDAGSELNPNQHVGVDEFPADLKFYWLHFVKAPDNDPRLWTHLS
VPOOTKVVDQALVALPROFLSBOEKHPTALEFVLVLIQOLTROARLDEQADGAG
TTLAWKAQLRIQYHRPLSTALLARELCYNADYLRVYRVEFLVLTALHQRVRE
AEKLLISEARLSLEVAALCGFNDVGYFRQIFRHTGLTPAARRYSKHINS"
1901..3301
/ gene="ECs4458"
1901..3301
/ gene="ECs4458"
/ note="probable permease, similar to hypothetical proteins [Salmonella typhimurium] gi|7442781|pir||C65167 percent identity 37 in 444 aa, also similar to transport proteins (putative symporters) e.g. YicJ [Escherichia coli (K-12)] gi|2851421|sp|P31435|YICJ_EC"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="putative permease"
/protein_id="BAB37881.1"
/db_xref="GI:13363933"
/translation="MTSTPTTTADIAKGIKDDALSVREKIGYGLGAGTVITCLLW
NFLTFFYDVRGLPALVGLFIALRVFDALSDPVMGVIADRTOSRMRGPRPQWLTAI
PIGIGILFTVPDASMGVIAWAGTYLLLSVGYTAIWPYCARINTWTRHNEVIS
CQSRFVLCGVAGFLVSGLVPWVAFPGQNAARGIOLGVGLCALVAVNFICCFFW
KERVPLSTMGKFTLRHLAGLRNDQLLMSFLINVNIRGGGYMTFTYVLOG
STGYLSFTFTWTFASIGSVISPLTRFDYK1XYTNNLLAALAMWFLPSPGA
YQTLWAVILGNGVILGFTPLRFLSLMAFADYGEWKTVRSGSMFAFLNFIKLA
ASSAGIISLFIFFVAVQPVENGTASSLGGIAAMETLLPALFHLLAGAIFRCKLNP
MMSRIATDLRQRVQ"
3312..5282
/ gene="ECs4459"
3312..5282
/ gene="ECs4459"
/ note="unknown, similar to hypothetical protein [thermotoga maritima] gi|7452109|pir||F72395 percent identity 37 in 635 aa"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="hypothetical protein"
/protein_id="BAB37882.1"
/db_xref="GI:13363934"
/translation="MNISEVDLHKLTVSDPFLGQYQQLVRDVVPIQWDALNDRIPEA
EPSHAIENFIAAGLGEFEYGMVQFSDVAKWLEAVNSLCQKPAELEKTADEVIE
LISAQCEDEGLYFTFYKAPERPWSNLAECHELICAGHLIEAGVAFQATGKRRLIG
VVCRLADHIDSFGPDSKLRGYPGHPH1ELALMRLYEVTERPRYLALTNFYFEQRA
PHYTDQVEYKRGQTSWHHTYGPAMWVKDKAYSOAHLPLAQOQTAIGHAVRFVYLMTG
VAHLARSHDSSKQDCRLRNWNAQRGLYITGGIGSSGSGEAFSSDYDLNDPVTYAE
SCASITGLMFAFMFARMLEMGDSOYADVMERALYNTVLGMDLGDGKHFYVNPLEVHPKS
LAFNHYLHVDPKPIRQWFGCACCPNIARVLTISGHLYTIPREDALYINIVAGSMEV
PVENGTLRLRVSGNPMQEOVTIAVESPOVPHRTIALRLPDMCTOPQIILAGEVEOD
IRKCYLHTREMGEGDTNLILPMPVRVYGNPLVRHVAGKATQRGPLVYCLEKADN
GESLHNLPLTDAFTTPEGKGLFSHKILIQAPGYEQSNPEQOQLWHYDSAPAKRO
TOTLTFIPWFSWANRGEGERIWNKEKHCP"
complement(5412..6152)
/ gene="ECs4460"
complement(5412..6152)
/ gene="ECs4460"
/ note="similar to YIAI_ECOLI gi|1790010 percent identity 99 in 246 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="putative outer membrane protein"
/protein_id="BAB37883.1"
/db_xref="GI:13363935"
/translation="MLINRNIVALFALPMSATASELSIGAGAAYNESPYRGYNENT
KAIPLSYEGDSFYVROTTLGPILSQSEKNLSLTASWMPLEFDPADNDYAMOODLK
RDSAMAGVAVYHREWGTVKASAADVLDSNGWVGLSVFHKQIGRLSILTPALGV
LYDENFSDYYIGISESSRRSLASYSNAQDAWVPVSVITAKYPIGEHVLMASAGYS
EUPEITDPSMIDRNESFTFTVGVSWRF"
6276..7250
/ gene="ECs4461"
6276..7250
/ gene="ECs4461"
/ note="similar to YIAI_ECOLI gi|1790011 percent identity 99 in 324 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="putative transcriptional regulator LYSR-type"
/protein_id="BAB37884.1"
/db_xref="GI:13363936"
/translation="MTKQLKYLRELKIIIVIAASENISHAATVGLIAQANVSKYLADF
ESKGLVKVDFDTRTROLLPFGTALLPYINDMLDRNOLNPFADYKHEKRGRTVIA
PGLITLYLSKHVIDIKDIDITLSLTCNLRNAPYEGVEFPDCCDLISVAPPKDE
SLVASFITQYAVTAYASORYLEKHPISRPDELEHSCILIDSMIDDIANIREFNVAGS
KEVDYRVKGNVCDNTQSALEARNHLGIVPAPKSKVSQSDLODGLTLPVCFQHPYEW
LDLVAIFRRKREYQWRVQVLDENLREINHLQALQSQLRPEQAESD"
complement(7247..8383)
/ gene="ECs4462"
complement(7247..8383)
/ gene="ECs4462"
/ note="similar to YIAI_ECOLI gi|1790012 percent identity 99 in 378 aa, similar to YIBH_ECOLI gi|1790024 percent identity 76 in 378 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence-not_experimental
/product="putative membrane protein"
/protein_id="BAB37885.1"
/db_xref="GI:13363937"
/translation="MDLLILTYVAFAMWFKIFKIPVKNKTIPTAALGGIFIVSGLI
LLMNVNHPYTFKAQKAVISIPVQVTVGVIEVTKKTLIKKCVLEFLDPTRYQAR
VRLNADIVTAHEKQALCAELDEMAANTQAKATRDFAKEYQRYAKSGSAKNPFS
EKDIDVARONIQAQSVKSSAAEQKIQSQSDLSVLVLSQIAQLAEAKYNLE"
Query Match 0.6%; Score 22; DB 1; Length 280900;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2452 atttatatcgttcgtcgaac 2473
|||||
DB 220079 ATTATATCCTGTCGTCAC 220100
|||||
RESULT 9
MMPOLE22
LOCUS MMPOLE22 824 bp DNA linear ROD 12-JAN-2000

```

```
DEFINITION Mus musculus DNA polymerase epsilon catalytic subunit (Pole) gene,
exons 48 and 49 and complete cds.
ACCESSION AF126398
VERSION AF126398.1
KEYWORDS GI:5669842
SEGMENT
SOURCE 22 of 22
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 824)
Huang, D., Knuuti, R., Palosaari, H., Pospiech, H. and Syvaoja, J. E.
cDNA and structural organization of the gene pole1 for the mouse
DNA polymerase epsilon catalytic subunit
Biochim. Biophys. Acta 1445 (3), 363-371 (1999)
99256371
10366722
2 (bases 1 to 824)
Huang, D., Knuuti, R., Palosaari, H., Pospiech, H. and Syvaoja, J.
Direct Submission
Submitted (08-FEB-1999) Department of Biochemistry, University of
Oulu, Linnanmaa, Oulu 90570, Finland
Location/Qualifiers
1..824
organism="Mus musculus"
strain="129/SvJ"
/db_xref="taxon:10090"
order(AF126377..1..1..2509,AF126378..1..1..7538,
AF126379..1..1..1157,AF126380..1..1..261,AF126381..1..1..697,
AF126382..1..1..234,AF126383..1..1..226,AF126384..1..1..553,
AF126385..1..1..800,AF126386..1..1..490,AF126387..1..1..295,
AF126388..1..1..501,AF126389..1..1..1340,AF126390..1..1..2714,
AF126391..1..1..273,AF126392..1..1..340,AF126393..1..1..311,
AF126394..1..1..288,AF126395..1..1..977,AF126396..1..1..273,
AF126397..1..1..403,1..824)
/gene="pole"
Join(AF126377..1..1728..1789,AF126378..1..893..1034,
AF126379..1..1442..1522,AF126378..1..1886..1930,
AF126378..1..2065..2157,AF126378..1..2835..2989,
AF126378..1..5125..5266,AF126378..1..5580..5660,
AF126378..1..6374..6481,AF126378..1..6757..6867,
AF126378..1..7093..7178,AF126378..1..7383..7502,
AF126379..1..51..183,AF126379..1..392..505,
AF126381..1..777..989,AF126380..1..51..158,AF126381..1..83..211,
AF126381..1..291..393,AF126381..1..480..626,
AF126382..1..167..212,AF126383..1..56..204,AF126384..1..84..176,
AF126384..1..363..507,AF126385..1..51..208,
AF126385..1..491..686,AF126386..1..28..242,
AF126387..1..100..202,AF126388..1..122..202,
AF126388..1..294..416,AF126389..1..247..456,
AF126389..1..683..892,AF126389..1..1073..1216,
AF126390..1..77..217,AF126390..1..469..622,
AF126390..1..699..805,AF126390..1..959..1135,
AF126390..1..1218..1441,AF126390..1..1622..1842,
AF126390..1..2050..2254,AF126391..1..37..210,
AF126392..1..99..224,AF126393..1..85..226,AF126394..1..53..230,
AF126392..1..199..224,AF126393..1..85..226,AF126394..1..53..230,
AF126395..1..97..228,AF126395..1..358..551,AF126396..1..46..246,
AF126397..1..130..255,133..222,307..630)
/gene="Pole"
/product="DNA polymerase epsilon catalytic subunit"
Join(AF126377..1..1728..1789,AF126378..1..893..1034,
AF126378..1..1442..1522,AF126378..1..1886..1930,
AF126378..1..2065..2157,AF126378..1..2835..2989,
AF126378..1..5125..5266,AF126378..1..5580..5660,
AF126378..1..6374..6481,AF126378..1..6757..6867,
AF126378..1..7093..7178,AF126378..1..7383..7502,
AF126379..1..51..183,AF126379..1..392..505,
AF126379..1..777..989,AF126380..1..51..158,AF126381..1..83..211,
AF126381..1..291..393,AF126381..1..480..626,
AF126382..1..167..212,AF126383..1..56..204,AF126384..1..84..176,
AF126384..1..363..507,AF126385..1..51..208,
AF126385..1..491..686,AF126386..1..28..242,
AF126387..1..100..202,AF126388..1..122..202,
AF126388..1..294..416,AF126389..1..247..456,
AF126389..1..683..892,AF126389..1..1073..1216,
AF126390..1..77..217,AF126390..1..469..622,
AF126390..1..699..805,AF126390..1..959..1135,
AF126390..1..1218..1441,AF126390..1..1622..1842,
AF126390..1..2050..2254,AF126391..1..37..210,
AF126392..1..99..224,AF126393..1..85..226,AF126394..1..53..230,
AF126392..1..199..224,AF126393..1..85..226,AF126394..1..53..230,
AF126395..1..97..228,AF126395..1..358..551,AF126396..1..46..246,
AF126397..1..130..255,133..222,307..630)
/gene="Pole"
/codon_start=1
/product="DNA polymerase epsilon catalytic subunit"
/db_xref="GI:5669844"
/translation="MVLNRNSRRHPPGADGKGRSDGSPSSVSALKRLERSQWTDM
DLRFGPERLKPGERTGNTLNNHPTIILDEDKRLVSADVYFIQDDGSKFKVALPYMP
YFYAARKGCDREVSSFLSKKKFOGKIAKLENPKEDLDLPNHLVGLKRSYIKLQSLNTI
EDLVKVKELSPAVKKNRODHASDEYTNLSSILOGGSVITDEDETSKIKADLONI
VDMREIDVPIHRLSIDLRHVAHWNTNFRGNAPVEITRRDDLVERDDPVLAFDI
ETKPLPTPDAETDQIMISYINIDCGGLITNREIVSEIDIEFETPRPEYGPFCY
FNPDEVLHIOQFHEIQETKPTIMVYNGDFDWPFEARAIHGLSQTIEIGFQD
SQGEYKAPQCIHMDCLRWKRSYLPVGSNKLAAAKLGDYDVELDPEDECMRMATE
QPOTLATYSVDATVYLYLMKYVHPFIIFALCTIIPMEPEVLKRSGLTCEALLAVQ
AFHANIIFPNKQSEENKLTDDGVLDAETVYGVHAEALSGVFRSDIICFRFENPAA
FDLIQVETKMRHAEIEEEKVPEQATNFQVEYEQIKTKLTSLKQWPNRCEPLIYH
LDVGAMPNIIILNRQPSAIVDEATCAACDNKPCAGCORMKAMQWGBFMPASRE
YHRIQHLESEKFPPLPEGPAPAFHLSREEOAKYKRLADYCKAKYKIHVTVKE
ERLTTTCORENSFYDVTAFRDBRYEFKGLHKWKKLSAAVEVGDASEVARKKME
ILDYDQAHKCLINSFPYGVYKRGARWMSMENAGIVCTGANIITQARELLIOIGRP
LEUDTQGLWCVLNPSFENPVINTNAKAPKLTISTPGALNINMVRGCTFNGHQSILT
EPKSLGLOIKIPOSSVFAFLKGLSTLEEVYGVAAKVKADYDLDVLYSKAAMPDSLF
VRSRGLYQIKIPOSSEVFAFLKGLSTLEEVYGVAAKVKADYDLDVLYSKAAMPDSLF
ELISENRSKLEIDYEQKSTSIKAKRLAEFLGDQWMDAGLSQRYIISRKPEESP
VTERAIPLAIFQAEPTRKHFRLKWLKSSLODFDITLDMDYIERGLSATQKIIIT
IPAAQOVNPRVPRKPDWLHKKLLEKNDYIKKKISLFLVLEGRKQIVMAQASNS
LSLCTPDMEDIGLTKPHSTVPVATKRWETQKESODIALTVPMQVGLGPPSLGT
TORENWLVQFHKKQLOAOORLARKKORLESADMPLGPIREGSTGLGSLFRK
TARSIMDLQWLIQISETQAGLFRWLWALIGNDLHCILSIPRVFYVNRVAKAEDGP
ALRYKNRNLPRNSIVNTLYEYSPEDMYTOHNEINTELSVPDIEGYETQVPLPRA
LVQGLGVYIPSRASVFLDVTYRNSOMFGLSALYSSEHLLDLDKVPKLLPPKHTF
HLKALGVYIPSRASVFLDVTYRNSOMFGLSALYSSEHLLDLDKVPKLLPPKHTF
EVAETNLETICRAIQFLLAYKEERGPTLJAVOSWELCRLTSEIPLERPLVPI
RVADKISYAVLDWRHGRARMIRHYLNDLCLISQAFEMSRYPHIPYCNLPEDISTGS
DELFARLHNNHLLWSPTSRPDLGCKEADNRLVMEFDNRATVINSGGYSTGVY
ELDIONLVNTIILQSHVNDMEGASNGISFDVIOQASLEDVMTGHOASALANYDET
ALGSSYFRLLAKSMVGVKATYHNIYADNOVMHFYRWLOSPCLLHDPALHRTLHN
MMKLLQLIAEKRLGSSVYVYANFNRIILCTKRRIEDALAVVEYITNLSKELFH
SLTISFSCWFEFLWNNNNILQFLQAASCSQYFLMIVSAIVAVYQSKBELRHS
EEEGMGSEVEDLENNNNILQFLQAASCSQYFLMIVSAIVAVYQSKBELRHS
PGSTPYKRGASQFSQESQATSLPGMITFSQDYVANELTOSFFITQKIOKKKVTGS
RNTTEPEMFPIPLGSHLLNPNLAFIRYKCVKLVSLDNTNITQVKNLRLDLVYD
GETSEEAQFDPCHSYVLPVEVICHSCNFCRDLCKDSSFSQDGAILPQWLKNCQAP
YDSSAIESALVEALQRLMAFTLQDLVCLKCGMKETHMPYVYCSGADFTLTIRTEVF
MEQIRIFONTAKYYSNSYLQETLEMLQTSVNSNC"
133..222
/gene="Pole"
/number=46
307..630
/gene="Pole"
/number=49
BASE COUNT 226 a 201 c 185 g 212 t
ORIGIN
exon
exon
Query Match 0.6% Score 21; DB 10; Length 824;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1401 aatcttcagaacattgccaa 1421
|||||
Db 327 AATCTTCAGAACATTGCCAA 347
```



```

RESULT 10
BC002275      1362 bp      mRNA      linear      ROD 12-JUL-2001
LOCUS
DEFINITION    Mus musculus, similar to DNA polymerase epsilon, clone MGC:7677
IMAGE:3496540, mRNA, complete cds.
ACCESSION    BC002275
VERSION      BC002275.1 GI:12805590
KEYWORDS     MGC.
SOURCE       house mouse.
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1362)
Strausberg,R.
Direct Submission
Submitted (31-JAN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 8 Row: h Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES             Location/Qualifiers
     source           1..1362
     /organism="Mus musculus"
     /db_xref="taxon:10090"
     /clone="MGC:7677 IMAGE:3496540"
     /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
     old, gross tissue."
     /clone_lib="NCL CGAP_Mam5"
     /lab_host="DH10B"
     /note="Vector: pCMV-SPORT6"
     /protein_id="AAH02275.1"
     /codon_start=1
     /product="Similar to DNA polymerase epsilon"
     /protein_id="AAH02275.1"
     /db_xref="GI:12805591"
     /translation="MTFQMSMSDHLSSHRAPSIILEYCPVFCDCPLPLSAAYIVAVYQS
     MKELRHAPGSTPYKRGASQFQSESGATGSLPGMITFSQDIYVANETQSFFTITQ
     KIQKVTGSRNTEPSEMFVLPGLSHLLNNPALEFIYVCKVSLDNTITQVKNLN
     RDLRLVDVGFSEBAQFRDPCHSVLPVETVCHSCNFCRDLDCDSSFSQDGAIIIPQ
     WLCNSQAPYPSSESALVEALORLMAFTLQDLVCLKRCGMKETHMPVYCSAGDF
     TLTRTFVFMQIRIFQNIAYKYSYSLQTEIEMLQYSPVSN"
BASE COUNT      353 a      363 c      298 g      348 t
ORIGIN

Query Match      0.6%; Score 21; DB 10; Length 1362;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 aatcttcagagaacattgccaa 1421
|||||
Db 1065 AATCTTCCAGACATTGCCAA 1085
|||||

RESULT 11
BC002275      2348 bp      mRNA      linear      VRT 06-JAN-2000
LOCUS
DEFINITION    Xenopus laevis prolactin receptor A mRNA, complete cds.
IMAGE:3496540, mRNA, complete cds.
ACCESSION    AF193800
VERSION      AF193800.1 GI:6180076
KEYWORDS     African clawed frog.
SOURCE       Xenopus laevis
ORGANISM     Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 2348)
Huang,H. and Brown,D.D.
Direct Submission
Submitted (12-OCT-1999) Embryology, Carnegie Institution of
Washington, 115 W. University Parkway, Baltimore, MD 21210, USA
Location/Qualifiers
     source           1..2348
     /organism="Xenopus laevis"
     /db_xref="taxon:8355"
     /note="PRLR-A"
     /codon_start=1
     /product="prolactin receptor A"
     /protein_id="AAF05776.1"
     /db_xref="GI:6180077"
     /translation="MPONLAPTPSVAVTFLFLYTVSLNQAQSLPGRVIDKRSYEK
     VTSCWMPGSDGLPTNYSLLYRKNDPKIYECPDVTVTSGINSCFFDKAHTSFWIF
     HIYVNFATNALGSNVSDESDVDTTVIVQYPPTNYSLVASEGHDLKLLKLPAPAVDVQ
     SGWLTLYEYRKQEWAEHLVGNQLKLFGLTPGGNYVQVRCKPDSGHWSVE
     STSYIQIPGGGRKTLDTLSIGALSAVICLTMTWTMLKRCILMSCILPLPVPQPKI
     MGDTQILLKSKSDELLSALGCGCFPTSDYDILLVEFVDDSKOHLIARSDSQOK
     QNIKVCPADTNDSDSGSCDSPFSHSGCKELRNQLQGVSDTRGDHHPQPTVMD
     GSFTNFSNKSNNMPPVPTSGNQISKSQYQIDKITNLIAISAMTASOHELLIPDNKN
     OLRCFKTEIIDEEKTKPSDFEDLRSGVEADTHLVLPNDKPPFLPARTMDYVEVHK
     VSQNNALALPKHNENSVRTQSVLVPNREYTKVERGNNVLVLQEMVNOVNPVT
     IESTKLKFAQQSANPNMFSLAQNGESRQTQAGGMYMDPSAFLS"
BASE COUNT      699 a      513 c      529 g      607 t
ORIGIN

Query Match      0.6%; Score 21; DB 5; Length 2348;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2200 cgggcagactttattattctg 2220
|||||
Db 1232 CGGCACACTTTTATATCTG 1212
|||||

RESULT 12
AB030443/c
LOCUS
DEFINITION    Xenopus laevis mRNA for prolactin receptor, complete cds.
IMAGE:3496540, mRNA, complete cds.
ACCESSION    AB030443
VERSION      AB030443.1 GI:6815120
KEYWORDS     prolactin receptor.
SOURCE       Xenopus laevis
ORGANISM     Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
1 (sites)
Yamamoto,T., Nakayama,Y., Tajima,T., Abe,S. and Kawahara,A.
Cloning of a cDNA for Xenopus prolactin receptor and its
metamorphic expression profile
Dev. Growth Differ. 42 (2), 167-174 (2000)
20289252

```

```
REFERENCE
AUTHORS Yamamoto, T.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1999) Takashi Yamamoto, Faculty of
Science, Kumamoto University, Department of Biological Science;
Kurokami 2-39-1, Kumamoto, Kumamoto 860-8555, Japan
(E-mail: tybig@po.kumamoto-u.ac.jp, Tel:81-96-342-3430(ex.3430),
Fax:81-96-342-3431)
FEATURES
source Location/Qualifiers
1..2879
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_lib="kidney"
208..2043
/codon_start=1
/product="prolactin receptor"
/protein_id="BAA90400.1"
/db_xref="GI:6815121"
/translatation="MPQNLAPTPPSVAVTFLFLYTVSLNAQLPKGPVTDKCRSEYK
VTFSCWKMKGSDGLPTNYSLLYKENDPKIYECPDYVTSGLNSCFDFDKAHTSFWIF
HIYVFNATNALGNSVDESDYTYIVQPYPTNVSLAVESCHHDLILKWLPPAMVDVQ
SCHWLTLYEVRYKEEKEQWEAHLVGNLKLKFLGTPGGNYVQVRCRKPDSGHWSW
STESYIQIFGGKKDLTLWISIGALSAVICLTIWMTALKRCSLASCILPPVGPKEI
MGDTQLLKSGLDELLSALGCGSPPTSDYDOLLVEFLVDSKQHLIASRQSQK
QNTKVCAPTNDNSGRSCDSPFSHSECKELNLPQVDSYDTRGIDHWPQNTVMD
QIFCTFNSDNKSNMPPVPTGNIKSQSYQIDITNLAISAMTASQHELLIPCNDKN
QLRCFKTIEIDETKTPSDFDLRSKGYEADTVHLVNDKPPFLPARMDVYEVHK
VSQNNALALIPKHNESVRTDOYSLVNPNEYTKVERVEGNNVLVLQEMVNVQNPVT
IESTKLKFOQSQSANPMNPSLAQNESRTQAGGMGYMDFSAFLS"
BASE COUNT 888 a 596 c 603 g 792 t
ORIGIN

Query Match 0.6%; Score 21; DB 5; Length 2879;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2200 cgggcagactttattctg 2220
|||||
DB 1230 CGGCGACACTTTTATTCTG 1210

RESULT 13
ARI49776
LOCUS ARI49776 4002 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6228615.
ACCESSION ARI49776
VERSION ARI49776.1 GI:15114367
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 4002)
AUTHORS Skatrud, P. Luther, de Waard, M.A., Andrade, A.C. and Peery, R. Brown.
TITLE Multiple drug resistance gene atrd of Aspergillus nidulans
JOURNAL Patent: US 6228615-A 1 08-MAY-2001;
FEATURES
source Location/Qualifiers
1..4002
/organism="unknown"
BASE COUNT 974 a 1034 c 1007 g 987 t
ORIGIN

Query Match 0.6%; Score 21; DB 6; Length 4002;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3553 gaagcagctcagccctcgac 3573
|||||
DB 3763 GAAGCGAGCTCAGCCCTCGAC 3783

RESULT 14
ARI49776
LOCUS ARI49776 4002 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6228615.
ACCESSION ARI49776
VERSION ARI49776.1 GI:15114368
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 4002)
AUTHORS Skatrud, P. Luther, de Waard, M.A., Andrade, A.C. and Peery, R. Brown.
TITLE Multiple drug resistance gene atrd of Aspergillus nidulans
JOURNAL Patent: US 6228615-A 3 08-MAY-2001;
FEATURES
source Location/Qualifiers
1..4002
/organism="unknown"
BASE COUNT 974 a 1034 c 1007 g 987 t
ORIGIN

Query Match 0.6%; Score 21; DB 6; Length 4002;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3553 gaagcagctcagccctcgac 3573
|||||
DB 3763 GAAGCGAGCTCAGCCCTCGAC 3783

RESULT 15
AF173826
LOCUS AF173826 4569 bp DNA linear PLN 03-FEB-2000
DEFINITION Emericella nidulans ABC-transporter (abcd) gene, complete cds.
ACCESSION AF173826
VERSION AF173826.1 GI:6856158
KEYWORDS
SOURCE
ORGANISM
Emericella nidulans.
REFERENCE
1 (bases 1 to 4569)
AUTHORS Nascimento, A.M., Terenzi, M.F., Goldman, M.H.S. and Goldman, G.H.
TITLE Molecular characterization of ABC-transporter encoding genes in
Aspergillus nidulans
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 4569)
AUTHORS Nascimento, A.M., Terenzi, M.F., Goldman, M.H.S. and Goldman, G.H.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1999) Ciencias Farmaceuticas, FCFRP-Universidade
de Sao Paulo, Av. do Cafe S/N, Ribeirao Preto, SP 14040-903, Brazil
FEATURES
source Location/Qualifiers
1..4569
/organism="Emericella nidulans"
/db_xref="taxon:162425"
/chromosome="VIII"
join(<1..405,457..3740,3798..>4155)
/gene="abcd"
/product="ABC-transporter"
<1..>4155
/gene="abcd"
join(1..405,457..3740,3798..4155)
/gene="abcd"
/codon_start=1
/product="ABC-transporter"
/protein_id="AAF29805.1"
/db_xref="GI:6856159"
/translatation="MSPLETNPLSPETAMREPATSTTEQASTTAEQASTHAADKKILSDLS
APSSTATPADKEHRPKSSSNVAVNDALIAHLIPEDERQVLTQLEEFKVNISF
FGLNRYATKMDILIMVISTICAIAAGALPLFTILFGLSLASTFORIMLYQISYDEFYD
ELKNVLYFVILGIGEPVTVYVSTVGFTTGERHATOKIREYLESILRONIGYFDKLG
AGEVTRITADNLIDQDISEKVGTLTALTATFATFIAYVYKWLKALICSSTVIAL
VLTMGGSQFIIKYSKSLDSGAGTVAEPISSIRNATATFGTQDKLAKOYEVHDE
AEKWGTKNOIVMGFMIGMFGMYSNYGLFNMGRFLVDGAVDVGDLITVLVAILIG
```

SFSLGNSPNAQFTNAVAAAQIFGTIDRQSPDYPYSEKGTLDHPEGHIELRNVKH  
IYPSRPEVTVMEDYSLMPAGKTTALVPGSGSKSTVVGGLVERFYMPVGTVLLDGH  
TKDLNRLWROQISLVSQEPVLFGTTIYKNIRHGLIGTKYENESDKVRELIENAKM  
ANADFTALPEGYETNVGORGFTLLSGGOKORIALARAVSDPKILLLDEATSA  
SEGVQALERAAREGRTTIVIAHRLSTIKTAHNIIVLVNGKIAEQGTHDELVRG  
KRLVQRINERKADALEDAEDLTNADIAKIKTASSASSDLGKPTTIDRTGTHK  
SVSAIUSKRPEPTPYSLWTLIKFVAFNRPPIPYMLIGLVFSLVAGGQPTQAVL  
YAKAISTLSLPESYSLRHDADFWLFFVVGIIQITQSTNGAAPAVGCSERLI  
RSTAFRTILRODIAFFDKENSTGALTSTETKHLGSGVGTGTLGTLMTSTLGA  
IIIALAIGWKALVCISVVPVLLACGYFYRLMAQFQSRKLAEGSANFACATSI  
RTVASLRERDVMEIYHAOLDQAQRTSLISVLRSSLLYASSQALVFCVALGFWYGT  
LLGHHEVDIPRFVCFSEILFQAOSACTVFSFAPDMGKAKNAAEFRRLFRKPOIDN  
WSEGEKLETFVEGEIERPNVHFYRTPREOPVLRLDLTVKPGYVALVPGSGGKST  
TIALLERFYDAIGSILVDGDKDISKLNINSYRSEFLSYSQEPTLYQGTIKENILLGIY  
EODVFEFLKACKADANIYDFIMSLPEGFNTVVGSGKMLSGGOKORVATARALLRDP  
KILLDEATSLADSEKVVQAAALDAARGTITIAVAHRLSTIQKADVIYVFDGKIV  
ESGTRSELVQKKGRYELVNLQSLGKH\*

BASE COUNT 1101 a 1186 c 1086 g 1196 t  
ORIGIN

Query Match 0.6%; Score 21; DB 8; Length 4569;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3553 gaagcagctcagccctcgac 3573  
          |||||  
Db 3913 GAAGCGACGTCAGCGCTCGAC 3933

Search completed: April 27, 2002, 05:55:01  
Job time: 18538 sec

